Differentiability of SDEs with drifts of super-linear growth

Citation for published version:
Imkeller, P, Reis, GD & Salkeld, W 2019, 'Differentiability of SDEs with drifts of super-linear growth', *Electronic journal of probability*, vol. 24, 3. https://doi.org/10.1214/18-EJP261

Digital Object Identifier (DOI):
10.1214/18-EJP261

Published In:
Electronic journal of probability

Differentiability of SDEs with drifts of super-linear growth

Peter Imkeller[1]
imkeller@mathematik.hu-berlin.de

Gonçalo Dos Reis[2,3] *
G.dosReis@ed.ac.uk

William Salkeld[2] †
w.j.salkeld@sms.ed.ac.uk

Greig Smith[3,4] ‡
G.Smith-13@sms.ed.ac.uk

[1]Humboldt Universität zu Berlin
[2]University of Edinburgh
[3]Centro de Matemática e Aplicações, (CMA), FCT, UNL
[4]MIGSAA

00h46, 20/03/2018

Abstract

We close an unexpected gap in the literature of stochastic differential equations (SDEs) with drifts of super linear growth (and random coefficients), namely, we prove Malliavin and Parametric Differentiability of such SDEs. The former is shown by proving Ray Absolute Continuity and Stochastic Gâteaux Differentiability. This method enables one to take limits in probability rather than mean square which bypasses the potentially non-integrable error terms from the unbounded drift. This issue is strongly linked with the difficulties of the standard methodology of [Nua06, Lemma 1.2.3] for this setting. Several examples illustrating the range and scope of our results are presented.

We close with parametric differentiability and recover representations linking both derivatives as well as a Bismut-Elworthy-Li formula.

Keywords: Malliavin Calculus, Parametric differentiability, monotone growth SDE, one-sided Lipschitz, Bismut-Elworthy-Li formula.

2010 AMS subject classifications: Primary: 60H07. Secondary: 60H10, 60H30

1 Introduction

In this manuscript we work with the class of Stochastic Differential Equations (SDEs) with drifts satisfying a super-linear growth (locally Lipschitz) and a monotonicity condition (also called one-sided Lipschitz condition). This class of SDEs appears ubiquitously in mathematics and engineering, for example, the stochastic Ginzburg-Landau equation in the theory of superconductivity; Stochastic Verhulst equation; Feller diffusion with logistic growth; Protein Kinetics and others, see [HJK11].

*G. dos Reis acknowledges support from the *Fundação para a Ciência e a Tecnologia* (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2013 (Centro de Matemática e Aplicações CMA/FCT/UNL).

†W. Salkeld acknowledges support from the Laura Wisewell travel fund

‡G. Smith was supported by The Maxwell Institute Graduate School in Analysis and its Applications, a Centre for Doctoral Training funded by the UK Engineering and Physical Sciences Research Council (grant [EP/L016508/01]), the Scottish Funding Council, Heriot-Watt University and the University of Edinburgh.

There is a wealth of results for the different differentiability properties of SDEs in general. However, and surprisingly, the landscape is (to the best of our knowledge) empty in respect to the *Monotone Growth* (or *One Sided Lipschitz*) setting apart from [TZ13] which we discuss below. Additionally, in [RS17] the authors discuss stochastic flows in rough path sense for a class related to ours.

In this work we establish Malliavin and parametric differentiability of this class of SDEs.

Malliavin differentiability. To establish Malliavin differentiability for an SDE with solution X and with monotone drifts the most natural path to follow is to try to apply [Nua06, Lemma 1.2.3] by employing a truncation procedure. The latter, yields a sequence X^n of SDEs with Lipschitz coefficients converging to X. Under said Lipschitz conditions the family X^n is Malliavin differentiable under suitable differentiability assumptions, with derivative DX^n, and one is able to appeal to [Nua06, Lemma 1.2.3] to conclude the Malliavin differentiability of X if one is able to show that $\sup_n \|DX^n\|_H < \infty$. The truncation procedure, even smoothed out, destroys the monotonicity and, in the multi-dimensional case, it is notoriously difficult to establish the mentioned uniform bound.

To the best of our knowledge this question was studied only in [TZ13]. There the authors employ a truncation procedure in order to use [Nua06, Lemma 1.2.3], unfortunately their [TZ13, Lemma 4.1] is incorrect. The constant M_l presented in their equation (4.1) depends on the truncation level n in a non-uniformly bounded way; the reader is invited to inspect the 2nd line of page 879. This lemma, which we were not able to fix, is used subsequently to establish the main result.

We prove Malliavin Differentiability through a less well-known method developed by Sugita [Sug85] which uses the concepts of *Ray absolute continuity* and *Stochastic Gâteaux Differentiability* see also [MPR17, IMPR16]. This approach is detailed in Section 3.2 below. The merit of this method is that the limit for the Stochastic Gâteaux derivative is a convergence in probability statement rather than a convergence in mean square statement. Put simply, this allows us to avoid cases such as the "Witches Hat" function where errors are non-integrable but converge to zero almost surely.

We study the case where the coefficients of the SDE are random. We follow the ideas of [GS16] and present two different sets of conditions which allow for Malliavin Differentiability. One set of conditions is sharp but somewhat difficult to prove. The other is much easier to verify but not sharp. We also provide examples discussing the scope and limitations of the approach.

Parametric differentiability. The second contribution of this work is parametric differentiability of such type of SDEs, and in particular, its implications for the classical case of deterministic coefficients. The methodology is inspired in that of the Malliavin differentiability section as we prove Gâteaux and Fréchet differentiability with respect to the SDEs parameters.

Representations, Absolute continuity of the law and Bismut-Elworthy-Li formulae. We bridge both differentiability results by recovering (a) representation formulae linking the Malliavin derivative and the parametric one; (b) establishing absolute continuity of the solution's Law; and (c) a Bismut-Elworthy-Li formula.

Technical results. In this setting the drift term is not bounded and conditional on the coefficients' integrability the solution may not be sufficiently integrable - see Remark 2.3 and the examples in Section 3.3. This means that the error terms appearing in proofs of differentiability will not be assumed to be sufficiently integrable. We negotiate this obstacle by proving everything in convergence in probability and ensuring that adequate conditions are met so that results can be extended to the relevant setting. Proposition 2.6 contains a Grönwall type inequality for the topology of Convergence in Probability that is of independent interest and is key to the methods used in this paper.

This paper is organized as follows. In Section 2, we lay out the notation and setting for this paper and recall a few baseline results from the literature. In Section 3 we prove Malliavin differentiability of SDEs of the form (2.1). There are two main results: Theorem 3.2 which provides a sharp method and Theorem 3.7 which has more comprehensive Assumptions but is not sharp. There is a collection of examples which explain the merits and limitations of the results we present.

In Section 4, we use similar methods to describe the Jacobian of the SDE. The final section bridges Section 2 and Section 4 and contains the so-called representations formulae and existence and smoothness results for densities.

2 Preliminaries

2.1 Notation and spaces

We denote by $\mathbb{N} = \{1, 2, \cdots\}$ the set of natural numbers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$; $\mathbb{R}$ denotes the set of real numbers respectively; $\mathbb{R}^+ = [0, \infty)$. By $a \lesssim b$ we denote the relation $a \leq C\, b$ where $C > 0$ is a generic constant independent of the relevant parameters and may take different values at each occurrence. By $\lfloor x \rfloor$ we denote the largest integer less than or equal to x. Let A be a $d \times m$ matrix, we denote the Transpose of A by A^T. When A is a matrix, we denote $|A|$ by $\mathrm{Tr}(A \cdot A^T)^{1/2}$.

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function. Then we denote ∇f to be the gradient operator and $H[f]$ to be the Hessian operator. ∂_{x_i} is the 1st partial derivative wrt i-th position. $\mathbb{1}_A$ denotes the usual indicator function over some set A

We use standard big O and little o notation to mean

$$f_n = O(f) \quad \Longleftrightarrow \quad \lim_{n\to\infty} \frac{f_n}{f} = C < \infty \qquad \text{and} \qquad f_n = o(f) \quad \Longleftrightarrow \quad \lim_{n\to\infty} \frac{f_n}{f} = 0.$$

where C is a constant independent of the limiting variable.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space carrying a m-dimensional Brownian Motion on the interval $[0, T]$; the Filtration on this space satisfies the usual assumptions. We denote by $\mathbb{E}$ and $\mathbb{E}[\cdot|\mathcal{F}_t]$ the usual expectation and conditional expectation operator (wrt to $\mathbb{P}$) respectively. For a random variable X we denote its probability distribution (or Law) by $\mathcal{L}^X$; the law of a process $(Y(t))_{t\in[0,T]}$ at time t is denoted by $\mathcal{L}^Y_t$.

Let $p \in [2, \infty)$. We introduce the following spaces and when there is no ambiguity about the underlying spaces or measures, we omit their arguments.

- Let $C([0, 1)$ denote the space of continuous functions $f : [0, 1] \to \mathbb{R}$ endowed with the uniform norm; $C_b([0, 1])$ its subspace of bounded functions; $C_b^k(\mathbb{R}^m)$ the set of k-times differentiable real valued maps defined on $\mathbb{R}^m$ with bounded partial derivatives up to order k, and $C_b^\infty(\mathbb{R}^m) = \cap_{k\geq 1} C_b^k(\mathbb{R}^m)$; C_b^0 its subspace of continuous bounded functions;
- Let $L^2([0, 1])$ denote the space of square integrable functions $f : [0, 1] \to \mathbb{R}$ satisfying $\|f\|_2 := \big(\int_0^1 |f(r)|^2 dr \big)^{1/2} < \infty$. Let H be the usual Cameron-Martin Hilbert space

$$H := \Big\{ h(t) = \int_0^t \dot{h}(s) ds,\ t \in [0, 1];\ h(0) = 0,\ \dot{h} \in L^2([0, 1]) \Big\}.$$

- Let $L^p(\mathcal{F}_t; \mathbb{R}^d; \mathbb{Q})$, $t \in [0, T]$, is the space of $\mathbb{R}^d$-valued $\mathcal{F}_t$-measurable RVs X with norm $\|X\|_{L^p} = \mathbb{E}^{\mathbb{Q}}[\,|X|^p]^{1/p} < \infty$; L^∞ refers to the subset of bounded RVs with norm $\|X\|_{L^\infty} = \operatorname{ess\,sup}_{\omega\in\Omega} |X(\omega)|$; Let $L^0(\mathcal{F}_t; \mathbb{R}^d)$ be the space of $\mathbb{R}^d$-valued $\mathcal{F}_t$ adapted random variables with the topology of convergence in measure.

- $\mathcal{S}^p([0,T],\mathbb{R}^m,\mathbb{Q})$ is the space of $\mathbb{R}^d$-valued measurable $\mathcal{F}$-adapted processes $(Y_t)_{t\in[0,T]}$ satisfying $\|Y\|_{\mathcal{S}^p} = \mathbb{E}^{\mathbb{Q}}[\sup_{t\in[0,T]}|Y(t)|^p]^{1/p} < \infty$; $\mathcal{S}^\infty$ refers to the subset of $\mathcal{S}^p([0,T],\mathbb{R}^m,\mathbb{Q})$ of absolutely uniformly bounded processes.
- $\mathbb{D}^{k,p}(\mathbb{R}^d)$ and $\mathbb{L}_{k,p}(\mathbb{R}^d)$ the spaces of Malliavin differentiable random variables and processes, see relevant section below.

2.2 Malliavin Calculus

Let $\mathcal{H}$ be a Hilbert space and $W:\mathcal{H}\to L^2(\Omega)$ a Gaussian random variable. The space $W(\mathcal{H})$ endowed with an inner product $\langle W(h_1),W(h_2)\rangle = \mathbb{E}[W(h_1)W(h_2)]$ is a Gaussian Hilbert space. Let $C_p^\infty(\mathbb{R}^n;\mathbb{R})$ be the space of all infinitely differentiable function which has all partial derivatives with polynomial growth. Let $\mathbb{S}$ be the collection of random variables $F:\Omega\to\mathbb{R}$ such that for $n\in\mathbb{N}$, $f\in C_p^\infty(\mathbb{R}^n;\mathbb{R})$ and $h_i\in\mathcal{H}$ can be written as $F=f(W(h_1),...,W(h_n))$. Then we define the derivative of F to be the $\mathcal{H}$ valued random variable

$$DF=\sum_{i=1}^n \partial_{x_i} f(W(h_1),...,W(h_n))h_i.$$

In the case of a stochastic integrals, $\mathcal{H}=L^2([0,T])$ and the Malliavin derivative takes the stochastic integral over a fixed interval to a stochastic process with parameter over the same fixed interval.

The Malliavin derivative from $L^p(\Omega)$ into $L^p(\Omega,\mathcal{H})$ is closable and the domain of the operator is defined to be $\mathbb{D}^{1,p}$. $\mathbb{D}^{1,p}$ is the closure of the of the set $\mathbb{S}$ with respect to the norm

$$\|F\|_{1,p}=\Big[\mathbb{E}[|F|^p]+\mathbb{E}[\|DF\|_{\mathcal{H}}^p]\Big]^{\frac{1}{p}}.$$

We also define the Directional Malliavin Derivative $D^hF=\langle DF,h\rangle$ for any choice of $h\in\mathcal{H}$. For more details, see [Nua06].

2.3 Existence and Uniqueness of SDE with Local Lipschitz coefficients

We present the class of SDEs with which we will be working.

Lipschitz and Local Lipschitz coefficients

Let $(t,\omega,\theta)\in[0,T]\times\Omega\times L^0(\Omega,\mathcal{F}_0,\mathbb{P};\mathbb{R}^d)$. In this paper, we prove differentiability properties of the SDE

$$X_\theta(t)(\omega)=\theta+\int_0^t b\Big(s,\omega,X(s)(\omega)\Big)ds+\int_0^t\sigma\Big(s,\omega,X(s)(\omega)\Big)dW(s), \tag{2.1}$$

driven by a m-dimensional Brownian motion W.

Assumption 2.1. *Let $p\geq 2$. Let $\theta:\Omega\to\mathbb{R}^d$, $b:[0,T]\times\Omega\times\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma:[0,T]\times\Omega\times\mathbb{R}^d\to\mathbb{R}^{d\times m}$ be progressively measurable maps and $L>0$ such that:*

- *$\theta\in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$ is independent of the Brownian motion W.*
- *b and σ are integrable in the sense that*

$$\mathbb{E}\Big[\Big(\int_0^T|b(t,\omega,0)|dt\Big)^p\Big],\mathbb{E}\Big[\Big(\int_0^T|\sigma(t,\omega,0)|^2dt\Big)^{\frac{p}{2}}\Big]<\infty. \tag{2.2}$$

- $\exists L$ *such that for almost all* $(s,\omega)\in[0,T]\times\Omega$ *and* $\forall x,y\in\mathbb{R}^d$ *we have*

$$\big\langle x-y, b(s,\omega,x)-b(s,\omega,x)\big\rangle_{\mathbb{R}^d}\leq L|x-y|^2 \quad \textit{and} \quad |\sigma(s,\omega,x)-\sigma(s,\omega,y)|\leq L|x-y|.$$

- *For* $x,y\in\mathbb{R}^d$ *such that* $|x|,|y|<N$ *and for almost all* $(s,\omega)\in[0,T]\times\Omega$*,* $\exists L_N>0$ *such that*

$$|b(s,\omega,x)-b(s,\omega,y)|\leq L_N|x-y|.$$

The next result extends other results found in the literature to the case of random coefficients. Existence and uniqueness of a solution follow the methods of [Mao08, Theorem 2.3.6]; the case of random coefficients is not addressed in his proof but his general methodology is applicable in the same way with only more care being taken when proving integrability.

Theorem 2.2. *Let* $p\geq 2$*. Suppose Assumption 2.1 is satisfied. Then there exists a unique solution* $(X(t))_{t\in[0,T]}$ *to the SDE* (2.1) *in* $\mathcal{S}^p$ *and*

$$\mathbb{E}\Big[\|X\|_\infty^p\Big]\lesssim\Big(\mathbb{E}\Big[|\theta|^p\Big]+\mathbb{E}\Big[\Big(\int_0^T|b(s,\omega,0)|ds\Big)^p\Big]+\mathbb{E}\Big[\Big(\int_0^T\Big|\sigma(s,\omega,0)\Big|^2ds\Big)^{\frac{p}{2}}\Big]\Big).$$

Moreover, the map $t\mapsto X(t)(\omega)$ *is* $\mathbb{P}$*-a.s. continuous.*

Finally, the solution of the SDE is Stochastically Stable in the sense that for $\forall\xi,\theta\in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$*,*

$$\mathbb{E}\Big[\|X_\xi-X_\theta\|_\infty^p\Big]\lesssim\mathbb{E}\Big[|\theta-\xi|^p\Big].$$

Proof. This proof can be found in Appendix A.1. ◻

Remark 2.3 (Issues with integrability and Fubini - Sharp conditions)**.** *The integrability conditions of Assumption 2.1 are designed to be sharp. However, they yield processes which can have some problematic properties.*

It is very important to note that we cannot (in general) swap the order of integration at this point! This is a key point in our manuscript. We are not able to assume that the drift term is sufficiently integrable (given (2.2)*) and hence the error terms appearing in the proofs of differentiability below will not be assumed to be integrable.*

To emphasize our point consider the following monotone drift function $b(t,\omega,x)=x-x^5$ *and* $\sigma(t,\omega,x)$ *is chosen so that for some* $t'\in[0,T)$

$$\mathbb{E}\Big[\Big(\int_0^T|\sigma(t,\omega,0)|^2dt\Big)^2\Big]<\infty,\quad\mathbb{E}\Big[\Big(\int_0^{t'}|\sigma(t,\omega,0)|^2dt\Big)^{\frac{5}{2}}\Big]=\infty.$$

These satisfy the conditions of Assumption 2.1 for $p=4$ *but not for* $p=5$*. We can then argue as follows: for* $t\in[t',T]$

$$\mathbb{E}[\,|X(t)|^4]<\infty,\quad\mathbb{E}[\,|X(t)|^5]=\infty.$$

The existence of finite fourth moments ensures we have finite first moments and hence for $t>t'$

$$\mathbb{E}\Big[\int_0^t\Big(X(s)-X(s)^5\Big)ds\Big]<\infty\quad\textit{which implies that}\quad\mathbb{E}\Big[\int_{t'}^tX(s)^5ds\Big]<\infty.$$

Linear Coefficients

Let $(t,\omega,\theta)\in[0,T]\times\Omega\times L^0(\Omega,\mathcal{F}_0,\mathbb{P};\mathbb{R}^d)$. We will also be interested in SDEs of the form

$$X_\theta(t)(\omega)=\theta+\int_0^t\Big[B\big(s,\omega\big)X_\theta(s)(\omega)+b(s,\omega)\Big]ds+\int_0^t\Big[\Sigma\big(s,\omega\big)X(s)(\omega)+\sigma(s,\omega)\Big]dW(s), \quad (2.3)$$

driven by a m-dimensional Brownian motion W. The derivatives of SDEs of the form (2.1) will satisfy linear SDEs of the form (2.3).

Assumption 2.4. *Let $p\geq 1$. Let $B:[0,T]\times\Omega\to\mathbb{R}^{d\times d}$, $\Sigma:[0,T]\times\Omega\to\mathbb{R}^{(d\times m)\times d}$, $b:[0,T]\times\Omega\to\mathbb{R}^d$ and $\sigma:[0,T]\times\Omega\to\mathbb{R}^{d\times m}$ be progressively measurable maps such that:*

- *$\theta\in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$ is independent of the Brownian motion W.*
- *B, b, Σ and σ are integrable in the sense that $\exists L\geq 0$ such that $\forall x\in\mathbb{R}^d$*

$$x^TB(t,\omega)x<L|x|^2\quad \mathbb{P}\text{-a.s.},\qquad \int_0^T\|\Sigma(t,\cdot)\|_{L^\infty}^2dt<\infty,$$
$$\mathbb{E}\Big[\Big(\int_0^T|b(t,\omega)|dt\Big)^p\Big],\quad \mathbb{E}\Big[\Big(\int_0^T|\sigma(t,\omega)|^2dt\Big)^{\frac{p}{2}}\Big]<\infty.$$

One advantage of SDEs of the form (2.3) is that they have an explicit solution unlike SDEs of the form(2.1) where a solution exists but cannot be explicitly stated. Linear SDEs do have Lipschitz coefficients, but their Lipschitz constants are not uniform over $(t,\omega)\in[0,T]\times\Omega$. Therefore, we cannot apply Theorem 2.2.

Notice that for Assumption 2.4, we do not make any requirement on B being positive definite operator. In fact, we may be interested in cases where $\exists x\in\mathbb{R}^d$ such that $x^T(\int_0^T B(t,\omega)dt)x=-\infty$ with positive probability.

Theorem 2.5. *Let $p\geq 1$. Suppose Assumption 2.4 is satisfied. Then there exists a unique solution $(X(t))_{t\in[0,T]}$ to the SDE (2.3) in $\mathcal{S}^p$ with explicit form*

$$X_\theta(t)=\Psi(t)\Bigg(\theta+\int_0^t\Psi(s)^{-1}\Big[b(s,\omega)-\Big\langle\Sigma(s,\omega),\sigma(s,\omega)\Big\rangle_{\mathbb{R}^m}\Big]ds+\int_0^t\Psi(s)^{-1}\sigma(s,\omega)dW(s)\Bigg),$$

where $\Psi:[0,T]\times\Omega\to\mathbb{R}^{d\times d}$ can be written as

$$\Psi(t)=I_d\exp\Bigg(\int_0^t\Big[B(s,\omega)-\frac{\big\langle\Sigma(s,\omega),\Sigma(s,\omega)\big\rangle_{\mathbb{R}^m}}{2}\Big]ds+\int_0^t\Sigma(s,\omega)dW(s)\Bigg), \quad (2.4)$$

and

$$\mathbb{E}\Big[\|X_\theta\|_\infty\Big]\lesssim\Big(\mathbb{E}\Big[|\theta|^p\Big]+\mathbb{E}\Big[\Big(\int_0^T|b(s,\omega)|ds\Big)^p\Big]+\mathbb{E}\Big[\Big(\int_0^T|\sigma(s,\omega)|^2ds\Big)^{\frac{p}{2}}\Big]\Big).$$

Moreover, the map $t\mapsto X(t)(\omega)$ is $\mathbb{P}$-a.s. continuous.

Finally, the solution of the equation is Stochastically stable in the sense that $\forall\xi,\theta\in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$

$$\mathbb{E}\Big[\|X_\xi-X_\theta\|_\infty^p\Big]\lesssim\mathbb{E}\Big[|\xi-\theta|^p\Big].$$

Proof. The proof of Existence and Uniqueness is found in [Mao08, Theorem 3.3.1]. Moment calculations are proved in Appendix A.1. Stochastic stability is proved in the same fashion as in Theorem 2.2. □

2.4 A Grönwall inequality

We introduce a Grönwall inequality in probability. To the best of our knowledge the next result is new and of independent interest. While unsurprising, this is key to the methods of this paper.

Proposition 2.6 (Grönwall Inequality for the Topology of Convergence in Probability). *Let $n \in \mathbb{N}$, $A_n : [0,T] \times \Omega \to \mathbb{R}$ be a sequence of adapted stochastic processes such that $\|A_n\|_\infty \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$.*

Let U_n be the solution of the SDE

$$U_n(t) = A_n(t) + \int_0^t f(U_n(s))ds + \int_0^t g(U_n(s))dW(s),$$

where $f, g : \mathbb{R} \to \mathbb{R}$ are Monotone growth and Lipschitz respectively (see 3rd bullet point of Assumption 2.1) and $f(0) = g(0) = 0$.

Then $\|U_n\|_\infty \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$.

Notice that since we do not have finite second moments of $\|A_n\|_\infty$, we cannot prove this using a mean square type argument.

Proof. Fix $\varepsilon' > 0$ and let $n \in \mathbb{N}$. We have that

$$\mathbb{P}\Big[\|U_n\|_\infty > \delta\Big] \leq \mathbb{P}\Big[\|U_n\|_\infty > \delta, \|A_n\|_\infty \leq \eta\Big] + \mathbb{P}\Big[\|A_n\|_\infty > \eta\Big],$$

for any choice of $\eta > 0$. We already have that $\lim_{n\to\infty} \mathbb{P}\big[\|A_n\|_\infty > \eta\big] = 0$ for any choice of $\eta > 0$ by assumption. Define the sequence of stopping times $\tau_n := \inf\{t' > 0 : |A_n(t')| > \eta\}$, $n \in \mathbb{N}$.

Firstly, we show that $\lim_{n\to\infty} \tau_n > T$ almost surely. Suppose this was not the case. The $\exists \Omega' \subset \Omega$ with $\mathbb{P}(\Omega') > 0$ and $\forall \omega \in \Omega'$ $\exists n_k(\omega)$ an increasing subsequence of integers such that $\tau_{n_k}(\omega) < T$ for all $k \in \mathbb{N}$. Then $\forall \omega \in \Omega'$, $\|A_{n_k}\|_\infty(\omega) > \eta$ for all $k \in \mathbb{N}$. But that implies that for any $k \in \mathbb{N}$ we have

$$\Omega' \subset \{\omega \in \Omega; \|A_{n_k}\|_\infty(\omega) > \eta\} \quad \text{and hence that} \quad \mathbb{P}\big[\|A_{n_k}\|_\infty > \eta\big] > \mathbb{P}[\Omega'].$$

The latter contradicts the assumption that $\|A_{n_k}\|_\infty$ converges to 0 in probability. So any such set Ω' must have measure 0 and we conclude $\lim_{n\to\infty} \tau_n > T$ almost surely.

The SDE for $U_n(t)$ is well defined for $t \in [0, \tau_n]$. Outside of this interval, A_n may not be integrable so we may not be able construct a solution. However $\forall \omega \in \Omega$ such that $\|A_n\|_\infty(\omega) \leq \eta$ we have that $\tau_n(\omega) > T$. Therefore

$$\mathbb{P}\Big[\|U_n(\cdot)\|_\infty > \delta, \|A_n\|_\infty \leq \eta\Big] = \mathbb{P}\Big[\|U_n(\cdot \wedge \tau_n)\|_\infty > \delta, \|A_n\|_\infty \leq \eta\Big]$$

because the process $U_n(\cdot)$ and the stopped process $U_n(\cdot \wedge \tau_n)$ are $\mathbb{P}$-almost surely equal when one restricts to the event where $\|A_n\|_\infty \leq \eta$.

As we know that the solution for $U_n(t \wedge \tau_n)$ will exist and make sense, it serves to introduce this stopping time. Thus we get

$$\begin{aligned}\mathbb{P}\Big[\|U_n\|_\infty > \delta\Big] &\leq \mathbb{P}\Big[\|U_n(\cdot \wedge \tau_n)\|_\infty > \delta, \|A_n\|_\infty \leq \eta\Big] + \mathbb{P}\Big[\|A_n\|_\infty > \eta\Big]\\ &\leq \mathbb{P}\Big[\|U_n(\cdot \wedge \tau_n)\|_\infty > \delta\Big] + \mathbb{P}\Big[\|A_n\|_\infty > \eta\Big].\end{aligned}$$

Now we consider the SDE for $U_n(t \wedge \tau_n)$. The stopping time prevents the term $A_n(t \wedge \tau_n)$ from getting any larger that η and ensures that the stochastic integral is a local martingale.

Hence we apply Theorem 2.2 to obtain existence/uniqueness of the solution and moment bounds.

$$\mathbb{E}\Big[\|U_n(\cdot\wedge\tau_n)\|_\infty^2\Big] < \eta^2 e^C \quad \text{and therefore} \quad \mathbb{P}\Big[\|U_n\|_\infty > \delta\Big] \le \frac{\eta^2 e^C}{\delta^2} + \mathbb{P}\Big[\|A_n\|_\infty > \eta\Big].$$

Choose η such that $\frac{\eta^2 e^C}{\delta^2} < \frac{\varepsilon'}{2}$. Then find $N \in \mathbb{N}$ such that $\forall n \ge N$ $\mathbb{P}\big[\|A_n\|_\infty > \eta\big] < \frac{\varepsilon'}{2}$. This concludes the proof. □

3 Malliavin Differentiability of SDEs with monotone coefficients

In this section we prove two Malliavin differentiability result for SDEs in the class given by Assumption 2.1. We use a less known method using the concepts of *Ray absolute continuity* and *Stochastic Gâteaux Differentiability* initiated by [Sug85] and later developed by [MPR17, IMPR16].

For SDE's of the form (2.1), the proof of existence and uniqueness of a solution involves a sequence of random variables which converge almost surely to the solution rather than in mean square. Indeed this sequence of random variables does not converge in mean square, unlike in the proof of Existence and Uniqueness for SDEs with Lipschitz coefficients. This means we are not able to employ the classical methods from [Nua06, Lemma 1.2.3] and recall our observation on the role that Proposition 2.6 will play here.

3.1 Main results and their assumptions

We state first the main assumptions/results with the proofs postponed to the later sections.

Assumption 3.1. *Let $b : [0,T]\times\Omega\times\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T]\times\Omega\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$ satisfy Assumption 2.1 for some $p > 2$. Further, suppose*

(i) For almost all $(t,\omega) \in [0,T]\times\Omega$ the functions $\sigma(t,\omega,\cdot)$ and $b(t,\omega,\cdot)$ have spatial partial derivatives in all directions.

(ii) For all $h \in H$ and $(\varepsilon, x) \in \mathbb{R}^+\times\mathbb{R}^d$, we have that the maps

$$(\varepsilon,x) \mapsto \int_0^T \Big|\nabla_x \sigma(t,\omega+\varepsilon h, x)\Big|^2 dt \qquad \text{and} \qquad (\varepsilon,x) \mapsto \int_0^T \Big|\nabla_x b(t,\omega+\varepsilon h, x)\Big|^2 dt,$$

are $\mathbb{P}$-a.s. jointly continuous.

(iii) $\exists U : [0,T]^2\times\Omega \to \mathbb{R}^{d\times m}$ and $V : [0,T]^2\times\Omega \to \mathbb{R}^{(d\times m)\times m}$ which satisfy that for $s > r$ $U(s,r,\omega) = V(s,r,\omega) = 0$ and

$$\mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^T \Big|U(s,r,\omega)\Big|^2 ds\Big)^{\frac{1}{2}} dr\Big)^p\Big] < \infty \quad \textit{and} \quad \mathbb{E}\Big[\Big(\int_0^T\int_0^T \Big|V(s,r,\omega)\Big|^2 ds dr\Big)^{\frac{p}{2}}\Big] < \infty.$$

(iv) b and σ satisfy, as $\varepsilon \to 0$, that $\forall h \in H$

$$\mathbb{E}\Big[\int_0^T \Big|\frac{b(r,\omega+\varepsilon h, X(r)) - b(r,\omega,X(r))}{\varepsilon} - \int_0^r U(s,r,\omega)\dot{h}(s) ds\Big| dr\Big] \to 0,$$
$$\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(r,\omega+\varepsilon h, X(r)) - \sigma(r,\omega,X(r))}{\varepsilon} - \int_0^r V(s,r,\omega)\dot{h}(s) ds\Big|^2 dr\Big] \to 0.$$

In the above condition neither b or σ are assumed to be in $\mathbb{D}^{1,2}$, they are only assumed to be Malliavin differentiable over the sub-manifold on which X (solution to (2.1)) takes values on. After our main results we give examples of SDE illustrating the scope of our assumptions.

Theorem 3.2 (Malliavin Differentiability of Monotone SDEs)**.** *Take $p > 2$. Let Assumption 3.1 hold and denote by X the unique solution of the SDE* (2.1) *in $\mathcal{S}^p$.*

Then X is Malliavin differentiable, i.e. $X \in \mathbb{D}^{1,p}(\mathcal{S}^p)$ and there exists adapted processes U and V such that the Malliavin derivative satisfies for $0 \leq s \leq t \leq T$

$$\begin{aligned} D_s X(t)(\omega) =& \sigma(s,\omega,X(s)(\omega)) + \int_s^t U(s,r,\omega)dr + \int_s^t V(s,r,\omega)dW(r) \\ &+ \int_s^t \nabla_x b(r,\omega,X(r)(\omega)) D_s X(r)(\omega)dr + \int_s^t \nabla_x \sigma(r,\omega,X(r)(\omega)) D_s X(r)(\omega)dW(r), \end{aligned} \tag{3.1}$$

and otherwise $D_s X(t) = 0$ for $s > t$.

The proof of Theorem 3.2 can be found in Section 3.4.

Remark 3.3 (Notation)**.** *At the simplest level, we have X is $\mathbb{R}^d$-valued and W is $\mathbb{R}^m$-valued. Therefore b, σ are $\mathbb{R}^d$- and $\mathbb{R}^{d\times m}$-valued respectively. Hence we have the collection of one-dimensional SDEs*

$$X^{(i)}(t)(\omega) = \theta^{(i)} + \int_0^t b^{(i)}(s,\omega,X(s)(\omega))ds + \sum_{j=1}^m \int_0^t \sigma^{(i,j)}(s,\omega,X(s)(\omega))dW^{(j)}(s),$$

where i is an integer between 1 and d.

The Malliavin Derivative DX is therefore an $\mathbb{R}^{d\times m}$ valued process and we get the system of equations

$$\begin{aligned} D_s^{(k)} X^{(i)}(t)(\omega) =& \sigma^{(i,k)}(s,\omega,X(s)(\omega))ds \\ &+ \int_s^t U^{(i,k)}(s,r,\omega)dr + \sum_{j=1}^m \int_s^t V^{(i,j,k)}(s,r,\omega)dW^{(j)}(r) \\ &+ \int_s^t \Big\langle (\nabla_x b^{(i)})(r,\omega,X(r)(\omega)), D_s^{(k)} X(t)(\omega) \Big\rangle_{\mathbb{R}^d} dr \\ &+ \sum_{j=1}^m \int_s^t \Big\langle (\nabla_x \sigma^{(i,j)})(r,\omega,X(r)(\omega)), D_s^{(k)} X(t)(\omega) \Big\rangle_{\mathbb{R}^d} dW^{(j)}(r), \end{aligned}$$

for i an integer between 1 and d and k an integer between 1 and m.

Remark 3.4 (Mollification and non-differentiability of b and σ)**.** *Using classic mollification arguments the assumptions of Theorem 3.2 concerning the behaviour of $x \mapsto b(\cdot,\cdot,x)$ and $x \mapsto \sigma(\cdot,\cdot,x)$ can be further weakened. Namely, σ can be assumed to be uniformly Lipschitz as opposed to continuously differentiable and b can be assumed to have left- and right-derivatives not necessarily equal to each other at every point.*

Under these conditions, a canonical mollification argument allows to re-obtain Theorem 3.2 where in (3.1) *one replaces $\nabla_x b$ and $\nabla_x \sigma$ by two processes corresponding to their generalized derivatives.*

If b and σ are assumed deterministic then one immediately obtains the familiar result.

Corollary 3.5 (Deterministic coefficients case)**.** *Suppose that $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ satisfy Assumption 2.1. Further, suppose that $x \mapsto b(\cdot, x)$ and $x \mapsto \sigma(\cdot, x)$ are continuously differentiable in their spatial variables (uniformly in t).*

Then X is Malliavin differentiable and for $0 \leq s \leq t \leq T$.

$$D_s X(t)(\omega) = \sigma(s, X(s)(\omega)) + \int_s^t \nabla_x b(r, \omega, X(r)(\omega)) D_s X(r)(\omega) dr \\ + \int_s^t \nabla_x \sigma(r, \omega, X(r)(\omega)) D_s X(r)(\omega) dW(r),$$

and otherwise $D_s X(t) = 0$ for $s > t$.

Assumption 3.1 can be strengthened to Assumption 3.6 which is much easier to verify.

Assumption 3.6. *Let $b : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ satisfy Assumption 2.1 for $p > 2$. Further, suppose Assumption 3.1 (i) and (ii) hold and*

(iii') b and σ are Malliavin differentiable in the sense that $\forall x \in \mathbb{R}^d$, $b(\cdot,\cdot,x) \in \mathbb{D}^{1,p}(L^1([0,T];\mathbb{R}^d))$ and $\sigma(\cdot,\cdot,x) \in \mathbb{D}^{1,p}(L^2([0,T];\mathbb{R}^{d\times m}))$,

(iv') The Malliavin derivatives of b and σ are Lipschitz in their spacial variables i.e. $\exists L > 0$ constant such that $\forall (s,t,\omega) \in [0,T]^2 \times \Omega$ and $x, y \in \mathbb{R}^d$, $\mathbb{P}$-almost surely

$$|D_s b(t,\omega,x) - D_s b(t,\omega,y)| \leq L|x-y|,$$
$$|D_s \sigma(t,\omega,x) - D_s \sigma(t,\omega,y)| \leq L|x-y|.$$

The second main result of the section is the following theorem.

Theorem 3.7. *Let $p > 2$. Let Assumption 2.1 hold and denote by X the unique solution of the SDE* (2.1) *in $\mathcal{S}^p$. Let b and σ satisfy Assumption 3.6. Then the conclusion of Theorem 3.2 still holds: $X \in \mathbb{D}^{1,p}(\mathcal{S}^p)$ and DX satisfies for $0 \leq s \leq t \leq T$*

$$\begin{aligned} D_s X(t)(\omega) =& \sigma(s,\omega, X(s)(\omega)) + \int_s^t (D_s b)(r,\omega,X(r)(\omega)) dr + \int_s^t (D_s \sigma)(r,\omega,X(r)(\omega)) dW(r) \\ &+ \int_s^t \nabla_x b(r,\omega,X(r)(\omega)) D_s X(r)(\omega) dr + \int_s^t \nabla_x \sigma(r,\omega,X(r)(\omega)) D_s X(r)(\omega) dW(r), \end{aligned} \tag{3.2}$$

and otherwise $D_s X(t) = 0$ for $s > t$.

The proof can be found in Section 3.5. We point out that the mollification Remark 3.4 applies to this result as well.

It is a well documented fact, see [Nua06, Theorem 2.2.1], that if one has an SDE with deterministic and Lipschitz drift and diffusion coefficients then the Malliavin derivative is the solution of a linear SDE. Both the SDE and the Malliavin Derivative have finite moments of all orders. Therefore the solution of the SDE exists in $\mathbb{D}^{1,\infty}$.

We study the case where the coefficients are random. SDEs of this kind do not always have finite moments of all orders, and the same will apply for the Malliavin derivative. In fact, the integrability of the derivative comes directly from the integrability of the Malliavin derivatives of b and σ.

3.2 Overview of the methodology and results on Wiener spaces

It is important to note that the solution of an SDE is not continuous with respect to $\omega \in \Omega$ (the *raison d'être* of rough paths). As the SDE exists in a probability space with the filtration generated by an m-dimensional Brownian motion, ω can be interpreted to mean the path of an individual Brownian motion plus any extra information about what happens when $t = 0$. However, it will be shown that the random variables are continuous, and indeed differentiable, when perturbed with respect to a path out of the Cameron Martin space. Hence for this section we take $h \in H^{\otimes m}$, an m-dimensional Cameron Martin path and $\dot{h}$ to be its derivative unless stated otherwise. We will not emphasize the difference between H and $H^{\otimes m}$ in this paper.

We start by introducing the concepts of *Ray absolute continuity* and *Stochastic Gâteaux Differentiability* and the results yielding Malliavin differentiability under those properties.

Let E be a separable metric space. Let $L(H, E)$ be the space of all bounded linear operators $V : H \to E$.

Definition 3.8 (Ray Absolutely Continuous map). *A measurable map $f : \Omega \to E$ is said to be* Ray Absolutely Continuous *if $\forall h \in H$, $\exists$ a measurable mapping $\tilde{f}_h : \Omega \to E$ such that*

$$\tilde{f}_h(\omega) = f(\omega) \quad \mathbb{P}\text{-a.e.}$$

and that $\forall \omega \in \Omega$,

$$t \mapsto \tilde{f}_h(\omega + th) \quad \textit{is absolutely continuous on any compact subset of } \mathbb{R}.$$

Definition 3.9 (Stochastically Gâteaux differentiable). *A measurable mapping $f : \Omega \to E$ is said to be* Stochastically Gâteaux differentiable *if there exists a measurable mapping $F : \Omega \to L(H, E)$ such that $\forall h \in H$,*

$$\frac{f(\omega + th) - f(\omega)}{t} \xrightarrow{\mathbb{P}} F(\omega)[h] \quad \textit{as } t \to 0.$$

Malliavin differentiability follows from [Sug85, Theorem 3.1] which was later improved upon by [MPR17, Theorem 4.1]. We recall both results next.

Theorem 3.10 ([Sug85]). *Let $p > 1$. The space $\mathbb{D}^{1,p}(E)$ is equivalent to the space of all random variables $f : \Omega \to E$ such that $f \in L^p(\Omega; E)$ is Ray Absolutely Continuous, Stochastically Gâteaux differentiable and the Stochastic Gâteaux derivative $F : \Omega \to L(H, E)$ is $F \in L^p(\Omega; L(H, E))$.*

Theorem 3.11 ([MPR17]). *Let $p > 1$. The space $\mathbb{D}^{1,p}(E)$ is equivalent to the space of all random variables $f : \Omega \to E$ such that $f \in L^p(\Omega; E)$ is Stochastically Gâteaux differentiable and the derivative converges in mean, e.g.*

$$\lim_{\varepsilon \to 0} \mathbb{E}\Big[\Big|\tfrac{f(\omega+\varepsilon h)-f(\omega)}{\varepsilon} - F(\omega)[h]\Big|\Big] \to 0. \tag{3.3}$$

This is called Strong Stochastic Gâteaux Differentiability.

The merit of [Sug85] is that it allows one to prove Malliavin differentiability by first establishing existence of a Gâteaux derivative and then extending to the full Frechét derivative. The convergence of the Gâteaux derivative in probability is a very weak condition that is much easier to prove than full Malliavin differentiability.

However, the requirement to additionally prove Ray Absolute Continuity makes this method more difficult. This is why [MPR17] extended this result to a stronger condition for convergence and removed the Ray Absolute Continuity condition.

Both of these methods have their merits. While studying different examples of processes with monotone growth, we became interested in the particular example where the drift term has polynomial growth of order q but only finite moments up to $p < q - 2$. In this case, one cannot in general find a dominating function for the error terms coming from the drift of the SDE while trying to prove Stochastic Gâteaux Differentiability. It therefore became necessary to prove only a convergence in probability statement, which required that we additionally prove Ray Absolute Continuity.

We state a useful result for the sequel, see [ÜZ00, Appendix B.1].

Proposition 3.12 (The Cameron-Martin Formula). *Let F be an $\mathcal{F}_T$-measurable random variable and let $h \in H$. Define $\mathcal{E}(h)(t) := \exp\Big(\int_0^t \dot{h}(s)dW(s) - \dfrac{1}{2}\int_0^t |\dot{h}(s)|^2 ds\Big)$ for $t \in [0,T]$.*

Then, when both sides are well defined,

$$\mathbb{E}\big[\, F(\omega + h) \,\big] = \mathbb{E}\Big[F \exp\Big(\int_0^T \dot{h}(s)dW(s) - \frac{1}{2}\int_0^T |\dot{h}(s)|^2 ds \Big)\Big] = \mathbb{E}\big[\, F(\omega)\mathcal{E}(h)(T) \,\big].$$

Moreover, $\forall h \in H$ and $\forall p \geq 1$ that $\mathcal{E}(h)(\cdot) \in \mathcal{S}^p([0,T])$.

Proof. The proof can be found in Appendix A.2. □

3.3 Examples

In this section, we discuss some interesting examples which emphasize the scope and sharpness of the assumptions made.

Example 3.13 (Concerning the continuity of $s \mapsto D_s X(\cdot)$). *Previous works on Malliavin calculus, see for example* [Nua06]*, treat the solution of this SDE as being continuous in s. While this is true for those examples studied, it is not true in the general case that we study here. We only have that it is $L^2([0,T])$ integrable. This example shows that it is not necessary for the derivative to be continuous in s. Take $g \in L^2([0,T])$ be a deterministic discontinuous function (a step function would be adequate) and assume the one dimensional setting. Consider σ of the form*

$$\sigma(t,\omega,x) = x + \int_0^t g(s)dW(s) \quad \textit{and} \quad b(t,\omega,x) = 0.$$

Hence $X(t)$ satisfies $X(t) = 1 + \int_0^t \big[X(s) + \int_0^s g(r)dW(r)\big]dW(s)$. It can be shown that the explicit solution of this equation is

$$X(t) = \exp\Big(W(t) - \frac{t}{2}\Big)\Big[1 - \int_0^t \int_0^r \exp\big(\frac{r}{2} - W(r)\big)g(p)dpdr + \int_0^t \int_0^r \exp\big(\frac{r}{2} - W(r)\big)g(p)dW(p)dr\Big].$$

Note that, as expected, X is a continuous process.

The process V, which represents the Malliavin derivative of σ, is

$$V(s,t,\omega) = D_s\sigma(t,\omega,X(t)(\omega)) = g(s)\mathbb{1}_{(0,t)}(s) \quad \Rightarrow \quad \int_s^t V(s,r,\omega)dW(r) = g(s)[W(t) - W(s)].$$

Clearly, the latter map is not continuous in s. The Malliavin derivative of X solves

$$D_sX(t) = X(s) + \int_0^s g(r)dW(r) + g(s)[W(t) - W(s)] + \int_s^t D_sX(r)dW(r).$$

Define $J_s(t) = \exp\Big([W(t) - W(s)] - \frac{t-s}{2}\Big)$*. Then the Malliavin derivative has the explicit solution*

$$D_s X(t) = J_s(t)\Big[X(s) + \int_0^s g(r)dW(r) + g(s)\Big(\int_s^t J_s(r)^{-1}dW(r) - \int_s^t J_s(r)^{-1}dr\Big)\Big].$$

Since g *is assumed to be not continuous, this will also not be continuous in* s*.*

We present a case where the coefficients are not Malliavin differentiable in general but are only differentiable on the set where the solution X takes its values. In other words, Assumption 3.1 is satisfied but Assumption 3.6 is not.

Example 3.14 (Malliavin Differentiable on the right manifold)**.** *Let* $d = m = 1$ *for simplicity. Let* $b(t, \omega, x) = -x$ *and*

$$\sigma(t, \omega, x) = \begin{cases} (x-1)^2(x+1)^2, & x \in [-1, 1] \\ \phi(x)f(\omega), & |x| > 1 \end{cases}$$

where $\phi \in C^\infty$*,* $\phi(x) = 0$ *for* $|x| \le 1$ *and* $\phi(x) = 1$ *for* $|x| \ge 2$*. The function* f *is any function* $f : \mathbb{R} \to \mathbb{R}$ *which is bounded, continuous but not differentiable and* ω *is the path of the Brownian motion.*

An example of such a function f *could be*

$$f(x) = \begin{cases} W'(x), & x \in [-1, 1] \\ -2, & |x| > 1 \end{cases}$$

where $W'(x)$ *is the Weierstrass function. The Weierstrass function is continuous but not differentiable anywhere and satisfies* $W'(-1) = W'(1) = -2$*. The latter implies that* f *is continuous. Hence* $f(W'(t))$ *will not be Malliavin differentiable but* $\varepsilon \mapsto f(W'(t) + \varepsilon h(t))$ *will be continuous.*

The derivative of σ *will satisfy*

$$\partial_x \sigma(\omega, x) = \begin{cases} 2x(x-1)(x+1), & x \in [-1, 1] \\ \phi'(x)f(\omega), & 1 < x < 2 \\ 0, & |x| > 2 \end{cases} \tag{3.4}$$

so since f *is bounded, we conclude that* σ *is Lipschitz* $\forall \omega \in \Omega$ *and differentiable.*

When the initial conditions determine that the process starts inside the interval $[-1, 1]$*, this is a so-called Wright-Fisher process (see* [MSS12]*) and the solution will remain within the interval* $[-1, 1]$ *with probability 1. This is important because the non-Malliavin Differentiability only affects the system when the process exits the* $[-1, 1]$ *interval. The conditions of Assumption 3.1 are satisfied but* $\sigma(\cdot, x)$ *is not Malliavin differentiable for all* $x \in \mathbb{R}^d$*.*

Remark 3.15 (The square-integrability case)**.** *In* [MPR17]*, it is proved that one can do away with the Ray Absolute Continuity condition if one can prove a* Strong Stochastic Gâteaux Differentiability *condition, see Theorem 3.11 and Equation* (3.3)*. However, in* [IMPR16]*, the authors provide a random variable* $Z \in \mathbb{D}^{1,2}$ *which is not Strong Stochastic Gâteaux differentiable in the sense that*

$$\mathbb{E}\Big[\Big|\frac{Z(\omega + \varepsilon h) - Z(\omega)}{\varepsilon} - D^h Z\Big|^2\Big] \nrightarrow 0, \quad \textit{as } \varepsilon \to 0.$$

It is however true that for all values $q \in [1, 2)$

$$\mathbb{E}\Big[\Big|\frac{Z(\omega + \varepsilon h) - Z(\omega)}{\varepsilon} - D^h Z\Big|^q\Big] \to 0, \quad \textit{as } \varepsilon \to 0.$$

In our framework, it is necessary to prove our convergence results in mean square due to the nature of the monotonicity property. Therefore we require that our SDE has finite moment of order p for some $p > 2$. However, in light of the example provided in [IMPR16]*, we believe (but do not show) that that there exists a case where the solution to an SDE of the form* (2.1) *which has finite moments of order up to $p = 2$ which is Malliavin Differentiable. Stochastic Gâteaux Differentiability would follow as before, but it was unclear to us how one would prove Ray Absolute Continuity of such a process.*

Remark 3.16 (The spatial Lipschitz condition for the Malliavin Derivatives of b and σ)**.** *In Assumption 3.6 (iv') we assume that Db and $D\sigma$ are Lipschitz in the spacial variable. We chose this condition because it is easy to verify and strong enough to ensure that $\forall x \in \mathbb{R}^d$*

$$\mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s b(t,\omega,X(t))|^2 ds\Big)^{\frac{1}{2}} dt\Big)^p\Big] < \infty, \quad \mathbb{E}\Big[\Big(\int_0^T \int_0^t |D_s \sigma(t,\omega,X(t))|^2 ds dt\Big)^{\frac{p}{2}}\Big] < \infty.$$

However, this condition is by no means necessary. One could consider the case where Db is locally Lipschitz in space and satisfies a linear growth condition and equivalently prove Theorem 3.7. However, the proof is more involved as it involves a careful interplay using Hölder's inequality between the maximal integrability of X, Db, $D\sigma$ and several other stochastic terms.

3.4 Proofs of the 1st main result - Theorem 3.2

In what follows, the choice of θ (the initial condition in (2.1)) does not affect the Malliavin derivative because θ is $\mathcal{F}_0$-measurable. If Y is $\mathcal{F}_t$-measurable then $D_s Y = 0$ for any $t < s$, see [Nua06, Corollary 1.2.1].

Existence and Uniqueness of the Malliavin derivative $D_s X(t)$

We start by establishing that (3.1) has a unique solution where X solves (2.1). At this point, nothing is said about the solution of (3.1) being the Malliavin derivative to X solution of (2.1).

Theorem 3.17. *Let $p > 2$. For $(s,t) \in [0,T]^2$, let X be the solution to the SDE* (2.1) *under Assumption 3.1. Let $(M_s(t))$ be defined by the matrix of $L^2([0,T])$-valued SDE*

$$\begin{aligned} M_s(t)(\omega) =& \sigma(s,\omega,X(s)(\omega)) + \int_s^t U(s,r,\omega)dr + \int_s^t V(s,r,\omega)dW(r) \\ &+ \int_s^t \nabla_x b(r,\omega,X(r)(\omega))M_s(r)(\omega)dr + \int_s^t \nabla_x \sigma(r,\omega,X(r)(\omega))M_s(r)(\omega)dW(r), \end{aligned} \tag{3.5}$$

for $s < t$ and $M_s(t) = 0$ for $s > t$.

Then a unique solution exists in $\mathcal{S}^p([0,T];L^2([0,T]))$ and the process has finite p^{th} moment so that the norm

$$\mathbb{E}\Big[\Big(\sup_{t\in[0,T]} \int_0^T |M_s(t)|^2 ds\Big)^{\frac{p}{2}}\Big] < \infty.$$

Observe that Equation (3.5) is linear in M, so the sharpness of the integrability is determined by the integrability of U, V and σ. In the trivial case where $U = V = 0$ and $\sigma = 1$ then M has finite moments of all orders.

Proof of Theorem 3.17. For brevity we omit the ω in the process and involved functions.

Equation (3.5) is not an SDE, but an Stochastic Partial Differential Equation (SPDE). Therefore, we need to extend results from Section 2. Let e_n be an orthonormal basis of the space $L^2([0,T];\mathbb{R}^m)$. This is a separable Hilbert space, so without loss of generality we can say the orthonormal basis is countably infinite. Let V_n be the linear span of the set $\{e_1,...,e_n\}$. Let $P_n : L^2([0,T];\mathbb{R}^m) \to V_n$ be the canonical projection operators

$$P_n[f](s) = \sum_{k=1}^{n} \langle f, e_k \rangle_{L^2([0,T];\mathbb{R}^m)} e_k(s).$$

Then it is clear that $\lim_{n\to\infty} \|P_n[f] - f\|_{L^2([0,T];\mathbb{R}^m)} = 0$. For $k \in \mathbb{N}$, consider the sequence of $\mathbb{R}^d$ valued SDEs

$$\begin{aligned} M_k(t) = &\int_0^t \sigma(s,\omega,X(s))e_k(s)ds + \int_0^t \Big(\int_0^r U(s,r,\omega)e_k(s)ds\Big)dr + \int_0^t \Big(\int_0^r V(s,r,\omega)e_k(s)ds\Big)dr \\ &+ \int_0^t \nabla_x b(r,\omega,X(r))M_k(r)dr + \int_0^t \nabla_x \sigma(r,\omega,X(r))M_k(r)dW(r). \end{aligned}$$

For each $k \in \mathbb{N}$ this is an example of a general d- dimensional Linear Stochastic differential equation of the form (2.3). Hence a unique solution exists for each k by Theorem 2.5.

Next define

$$M_{(n),s}(t) = \sum_{k=1}^{n} M_k(t) e_k(s)^T \mathbb{1}_{[0,t)}(s),$$

which satisfies the SDE

$$\begin{aligned} M_{(n),s}(t) =&P_n\Big[\sigma(\cdot,\omega,X(\cdot))\Big](s) + \int_s^t P_n\Big[U(\cdot,r,\omega)\Big](s)dr + \int_s^t P_n\Big[V(\cdot,r,\omega)\Big](s)dW(r) \\ &+ \int_s^t \nabla_x b(r,\omega,X(r)(\omega))M_{(n),s}(r)(\omega)dr + \int_s^t \nabla_x\sigma(r,\omega,X(r)(\omega))M_{(n),s}(r)(\omega)dW(r). \end{aligned}$$

This SDE makes sense because the projection space is finite dimensional so we can rewrite it in a finite dimensional vector form. Finally, we take a limit as $n \to \infty$ and we see the coefficients of the SDE will converge to the coefficients of (3.5). To prove that the limit of the solutions converges, we take the norm of M in $\mathcal{S}^p$ to show it is finite then use the Dominated Convergence Theorem

The Malliavin Derivative DX is a matrix valued process. Therefore we require a norm on the space of L^2-valued matrices. Let $a^{(i,j)} \in L^2([0,T])$ and $A(s) = (a^{(i,j)}(s))_{i\in\{1,...,d\},j\in\{1,...,m\}}$. Then define

$$\|A.\| = \Big(\sum_{i=1}^{d}\sum_{j=1}^{m}\int_0^T |a^{(i,j)}(s)|^2 ds\Big)^{1/2}.$$

Firstly, we are going to apply Itô's formula for $f(x) = |x|^2$ while fixing $s \in [0, t]$. This yields

$$\begin{aligned}|M_s^{(i,k)}(t)|^2 =&|\sigma^{(i,k)}(s, X(s))|^2 \\ &+ 2\int_s^t M_s^{(i,k)}(r)\Big[U^{(i,k)}(s,r) + \Big\langle \nabla_x b^{(i)}(r, X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\Big]dr \\ &+ 2\sum_{j=1}^m \int_s^t M_s^{(i,k)}(r)\Big[V^{(i,j,k)}(s,r) + \Big\langle \nabla_x \sigma^{(i,j)}(r, X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\Big]dW^{(j)}(r) \\ &+ \sum_{j=1}^m \int_s^t \Big[V^{(i,j,k)}(s,r) + \Big\langle \nabla_x \sigma^{(i,j)}(r, X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\Big]^2 dr.\end{aligned}$$

Integrating over s and since every term is positive, we can change the order of integration to obtain

$$\begin{aligned}&\int_0^t |M_s^{(i,k)}(t)|^2 ds = \int_0^t |\sigma^{(i,k)}(s, X(s))|^2 ds \\ &+ 2\int_0^t \int_0^r M_s^{(i,k)}(r)\Big[U^{(i,k)}(s,r) + \Big\langle \nabla_x b^{(i)}(r, X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\Big]dsdr \\ &+ 2\sum_{j=1}^m \int_0^t \int_0^r M_s^{(i,k)}(r)\Big[V^{(i,j,k)}(s,r) + \Big\langle \nabla_x \sigma^{(i,j)}(r, X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\Big]dsdW^{(j)}(r) \\ &+ \sum_{j=1}^m \int_0^t \int_0^r \Big[V^{(i,j,k)}(s,r) + \Big\langle \nabla_x \sigma^{(i,k)}(r, X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\Big]^2 dsdr.\end{aligned}$$

We use Itô's formula $g(x) = \big(\sum_i x^{(i)}\big)^{p/2}$,

$$\|M.(t)\|^p = \Big(\sum_{i,k}\int_0^t |M_s^{(i,k)}(t)|^2 ds\Big)^{\frac{p}{2}} = \Big(\sum_{i,k}\int_0^t |\sigma^{(i,k)}(s, X(s))|^2 ds\Big)^{\frac{p}{2}}$$

$$+ p\int_0^t \|M.(r)\|^{p-2}\Big(\sum_{i,k}\int_0^r M_s^{(i,k)}(r)\Big[U^{(i,k)}(s,r) + \Big\langle \nabla_x b^{(i)}(r, X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\Big]ds\Big)dr \tag{3.6}$$

$$+ \frac{p}{2}\sum_{i,j,k}\int_0^t \|M.(r)\|^{p-2}\Big(\int_0^r \Big[V^{(i,j,k)}(s,r) + \Big\langle \nabla_x \sigma^{(i,j)}(r, X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\Big]^2 ds\Big)dr \tag{3.7}$$

$$+ p\sum_{i,j,k}\int_0^t \|M.(r)\|^{p-2}\Big(\int_0^r M_s^{(i,k)}(r)\Big[V^{(i,j,k)}(s,r) + \Big\langle \nabla_x \sigma^{(i,j)}(r, X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\Big]ds\Big)dW^{(j)}(r) \tag{3.8}$$

$$+ p(p-2)\int_0^t \|M.(r)\|^{p-4}\sum_{i,j,k}\Big(\int_0^r M_s^{(i,k)}(r)\Big[V^{(i,j,k)}(s,r) + \Big\langle \nabla_x \sigma^{(i,j)}(r, X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\Big]ds\Big)^2 dr. \tag{3.9}$$

We take a supremum over $t \in [0, T]$ then expectations and we find an upper bound for $\mathbb{E}[\,\|M(\cdot)\|_\infty^p]$.

Using Assumption 3.1, we have that

$$\mathbb{E}\Big[\Big(\int_0^T \|V(\cdot, r)\|_2^2 dr\Big)^{\frac{p}{2}}\Big] < \infty, \qquad \text{and} \qquad \mathbb{E}\Big[\Big(\int_0^T \|U(\cdot, r)\|_2 dr\Big)^{p}\Big] < \infty.$$

Let $n \in \mathbb{N}$ be an integer which we will choose later. We consider each of these 4 terms above. Firstly,

$$\text{For (3.6)} \Rightarrow p\mathbb{E}\Big[\int_0^T \|M.(r)\|^{p-2}\Big(\sum_{i,k}\int_0^r M_s^{(i,k)}(r)U^{(i,k)}(s,r)ds\Big)dr\Big] \tag{3.10}$$

$$+ p\mathbb{E}\Big[\int_0^T \|M.(r)\|^{p-2}\Big(\sum_{i,k}\int_0^r M_s^{(i,k)}(r)\Big\langle\nabla_x b^{(i)}(r,X(r)), M_s^{(\cdot,k)}(r)\Big\rangle ds\Big)dr\Big]. \tag{3.11}$$

Now

$$\begin{aligned}
(3.10) \leq & p\mathbb{E}\Big[\|M.\|_\infty^{p-1}\int_0^T \Big(\sum_{i,k}\int_0^r |U^{(i,k)}(s,r)|^2 ds\Big)^{\frac{1}{2}}dr\Big] \\
\leq & \frac{\mathbb{E}\Big[\|M.\|_\infty^p\Big]}{n} + [n(p-1)]^{p-1}\mathbb{E}\Big[\Big(\int_0^T \|U(\cdot,r)\|_2 dr\Big)^p\Big].
\end{aligned}$$

and $(3.11) \leq pL\int_0^T \mathbb{E}\Big[\|M.(r)\|^p\Big]dr$ using the monotonicity properties of b. Secondly,

$$\begin{aligned}
\text{for (3.7)} \Rightarrow & p\mathbb{E}\Big[\int_0^T \|M.(r)\|^{p-2}\Big(\sum_{i,j,k}\int_0^r |V^{(i,j,k)}(s,r)|^2 ds\Big)dr\Big] \\
& + p\mathbb{E}\Big[\int_0^T \|M.(r)\|^{p-2}\Big(\sum_{i,j,k}\int_0^r \Big\langle\nabla_x\sigma^{(i,j)}(r,X(r)), M_s^{(\cdot,k)}(r)\Big\rangle^2 ds\Big)dr\Big] \\
\leq & \frac{\mathbb{E}\Big[\|M.\|_\infty^p\Big]}{n} + 2[n(p-2)]^{\frac{p-2}{2}}\mathbb{E}\Big[\Big(\int_0^T \|V(\cdot,r)\|_2^2 dr\Big)^{\frac{p}{2}}\Big] + pL^2\int_0^t \mathbb{E}\Big[\|M.(r)\|^p\Big]dr.
\end{aligned}$$

Thirdly, using the Burkholder-Davis-Gundy Inequality

$$\begin{aligned}
\text{for (3.8)} \Rightarrow & pC_1\mathbb{E}\Big[\Big(\int_0^T \|M.(r)\|^{2p-4}\sum_j\Big(\sum_{i,k}\int_0^r M_s^{(i,k)}(r)\big[V^{(i,j,k)}(s,r) \\
& + \Big\langle\nabla_x\sigma^{(i,j)}(r,X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\big]ds\Big)^2 dr\Big)^{\frac{1}{2}}\Big]
\end{aligned}$$

$$\leq \sqrt{2}pC_1\mathbb{E}\Big[\|M.\|_\infty^{p-2}\Big(\int_0^T \Big[\sum_{i,j,k}\int_0^r |M_s^{(i,k)}(r)|\cdot|V^{(i,j,k)}(s,r)|ds\Big]^2 dr\Big)^{\frac{1}{2}}\Big] \tag{3.12}$$

$$+ \sqrt{2}pC_1\mathbb{E}\Big[\|M.\|_\infty^{p-2}\Big(\int_0^T \Big[\sum_{i,j,k}\int_0^r |M_s^{(i,k)}(r)|\cdot\Big|\Big\langle\nabla_x\sigma^{(i,j)}(r,X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\Big|ds\Big]^2 dr\Big)^{\frac{1}{2}}\Big]. \tag{3.13}$$

As before, we have

$$\begin{aligned}
(3.12) \leq & \sqrt{2}pC_1\mathbb{E}\Big[\|M.\|_\infty^{p-1}\Big(\int_0^T \|V(\cdot,r)\|_2^2 dr\Big)^{\frac{1}{2}}\Big] \\
\leq & \frac{\mathbb{E}\Big[\|M.\|_\infty^p\Big]}{n} + (\sqrt{2}C_1)^p[n(p-1)]^{p-1}\mathbb{E}\Big[\Big(\int_0^T \|V(\cdot,r)\|_2^2 dr\Big)^{\frac{p}{2}}\Big].
\end{aligned}$$

and put together

$$(3.13) \leq \sqrt{2} p C_1 L \mathbb{E}\Big[\Big(\int_0^T \|M.(r)\|^{2p} dr\Big)^{\frac{1}{2}}\Big] \leq \sqrt{2} p C_1 L \mathbb{E}\Big[\|M.\|_\infty^{\frac{p}{2}}\Big(\int_0^T \|M.(r)\|^p dr\Big)^{\frac{1}{2}}\Big]$$
$$\leq \frac{\mathbb{E}\Big[\|M.\|_\infty^p\Big]}{n} + \frac{(pC_1L)^2 n}{2}\int_0^T \mathbb{E}\Big[\|M.\|_{\infty,r}^p\Big] dr.$$

Finally, for

$$(3.9) \Rightarrow 2p(p-2)\mathbb{E}\Big[\int_0^T \|M.(r)\|^{p-4}\Big(\sum_{i,j,k}\int_0^r |M_s^{(i,k)}(r)| \cdot |V^{(i,j,k)}(s,r)| ds\Big)^2 dr\Big] \tag{3.14}$$

$$+ 2p(p-2)\mathbb{E}\Big[\int_0^T \|M.(r)\|^{p-4}\Big(\sum_{i,j,k}\int_0^r |M_s^{(i,k)}(r)| \cdot \Big|\Big\langle \nabla_x \sigma^{(i,j)}(r, X(r)), M_s^{(\cdot,k)}(r)\Big\rangle\Big| ds\Big)^2 dr\Big]. \tag{3.15}$$

Repeating the same ideas as before, we get

$$(3.14) \leq 2p(p-2)\mathbb{E}\Big[\|M.\|_\infty^{p-2} \cdot \int_0^T \|V(\cdot,r)\|_2^2 dr\Big]$$
$$\leq \frac{\mathbb{E}\Big[\|M.\|_\infty^p\Big]}{n} + 2^{\frac{p+2}{2}} \cdot n^{\frac{p-2}{2}} \cdot (p-2)^{p-1}\mathbb{E}\Big[\Big(\int_0^T \|V(\cdot,r)\|_2^2 dr\Big)^{\frac{p}{2}}\Big].$$

and $(3.15) \leq 2p(p-2)L^2\mathbb{E}\big[\int_0^T \|M.\|_{\infty,r}^p dr\big]$. Therefore, choosing $n = 6$ we conclude

$$\frac{\mathbb{E}\Big[\|M.\|_\infty^p\Big]}{6} \leq \Bigg(\mathbb{E}\Big[\|\sigma(\cdot, X(\cdot))\|_2^p\Big] + \widetilde{C}_1\mathbb{E}\Big[\Big(\int_0^T \|U(\cdot,r)\|_2 dr\Big)^p\Big]$$
$$+ \widetilde{C}_2\mathbb{E}\Big[\Big(\int_0^T \|V(\cdot,r)\|_2^2 dr\Big)^{\frac{p}{2}}\Big]\Bigg) + \widetilde{C}_3\int_0^T \mathbb{E}\Big[\|M.\|_{\infty,r}^p\Big] dr.$$

By applying Gronwall's inequality we conclude that $\mathbb{E}[\,\|M.\|_\infty^p] < \infty$. □

Ray Absolute Continuity of X

We show that the expectation of $\|\big(X(\cdot)(\omega+\varepsilon h) - X(\cdot)(\omega)\big)/\varepsilon\|_\infty^2$ has a bound uniform in ε. This relies on having finite p^{th} moments of the random variable $\|X\|_\infty$ for $p > 2$. If we only have finite second moments, this would not be true in general.

Proposition 3.18. *Let X be solution to the SDE* (2.1) *under Assumption 3.1. We have*

$$\mathbb{E}\Big[\Big\|\tfrac{X(\cdot)(\omega+\varepsilon h) - X(\cdot)(\omega)}{\varepsilon}\Big\|_\infty^2\Big] = O(1) \quad \textit{as } \varepsilon \to 0. \tag{3.16}$$

For any ε such that $0 < \varepsilon < 1$ the random variable

$$\frac{1}{\varepsilon}\|X(\cdot)(\omega+\varepsilon h) - X(\cdot)(\omega)\|_\infty < N(\omega) \quad \mathbb{P}\textit{-almost surely.} \tag{3.17}$$

where N is a square integrable random variable independent of ε. Also

$$\lim_{\varepsilon\to 0}\|X(\cdot)(\omega+\varepsilon h) - X(\cdot)(\omega)\|_\infty = 0.$$

Proof. Using Assumption 3.1, we can conclude that, as $\varepsilon \to 0$,

$$\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega,X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big] = O(1),$$

$$\mathbb{E}\Big[\Big(\int_0^T \Big|\frac{b(s,\omega+\varepsilon h, X(s)(\omega)) - b(s,\omega,X(s)(\omega))}{\varepsilon}\Big| ds\Big)^2\Big] = O(1).$$

We have

$$\begin{aligned}\frac{X(t)(\omega+\varepsilon h) - X(t)(\omega)}{\varepsilon} =& \int_0^t \sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h))\dot{h}(s)ds\\
&+ \int_0^t \Big(\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega,X(s)(\omega))\Big)\dot{h}(s)ds\\
&+ \frac{1}{\varepsilon}\int_0^t \Big(b(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - b(s,\omega,X(s)(\omega))\Big)ds\\
&+ \frac{1}{\varepsilon}\int_0^t \Big(\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega,X(s)(\omega))\Big)dW(s).\end{aligned}$$

Using Itô's formula for $f(x)=x^2$ and writing $P(t) = \frac{X(t)(\omega+\varepsilon h)-X(t)(\omega)}{\varepsilon}$ we have

$$\Big|P(t)\Big|^2 = 2\int_0^t \Big\langle P(s), \sigma(s,\omega,X(s)(\omega))\dot{h}(s)\Big\rangle ds \tag{3.18}$$

$$+2\int_0^t \Big\langle P(s), \Big(\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h,X(s)(\omega))\Big)\dot{h}(s)\Big\rangle ds \tag{3.19}$$

$$+2\int_0^t \Big\langle P(s), \Big(\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega,X(s)(\omega))\Big)\dot{h}(s)\Big\rangle ds \tag{3.20}$$

$$+2\int_0^t \Big\langle P(s), \frac{b(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - b(s,\omega+\varepsilon h,X(s)(\omega))}{\varepsilon}\Big\rangle ds \tag{3.21}$$

$$+2\int_0^t \Big\langle P(s), \frac{b(s,\omega+\varepsilon h, X(s)(\omega)) - b(s,\omega,X(s)(\omega))}{\varepsilon}\Big\rangle ds \tag{3.22}$$

$$+2\int_0^t \Big\langle P(s), \frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h,X(s)(\omega))}{\varepsilon}dW(s)\Big\rangle \tag{3.23}$$

$$+2\int_0^t \Big\langle P(s), \frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega,X(s)(\omega))}{\varepsilon}dW(s)\Big\rangle \tag{3.24}$$

$$+\int_0^t \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega,X(s)(\omega))}{\varepsilon}\Big|^2 ds. \tag{3.25}$$

We take a supremum over t then expectations. Let n be an integer that we will choose later. By using a combination of Young's Inequality, Cauchy-Schwartz Inequality and Burkholder-Davis-Gundy Inequality and the continuity properties from Assumption 3.1 allows us to find the following upper bounds:

$$\begin{aligned}\text{For (3.18)} \Rightarrow& \mathbb{E}\Big[2\int_0^T \Big|\Big\langle P(s), \sigma(s,\omega,X(s)(\omega))\dot{h}(s)\Big\rangle\Big| ds\Big]\\
\le& \frac{\mathbb{E}[\|P\|_\infty^2]}{n} + n\|\dot{h}\|_2^2\mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega,X(s)(\omega))\Big|^2 ds\Big]\\
\le& \frac{\mathbb{E}[\|P\|_\infty^2]}{n} + 2n\|\dot{h}\|_2^2\Big(L^2\mathbb{E}\Big[\|X\|_\infty^2\Big] + \mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega,0)\Big|^2 ds\Big]\Big),\end{aligned}$$

$$\text{For (3.19)} \Rightarrow \mathbb{E}\Big[2\int_0^T \Big|\Big\langle P(s), \Big(\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X(s)(\omega))\Big)\dot{h}(s)\Big\rangle\Big| ds\Big]$$
$$\leq 2L\varepsilon \int_0^T \mathbb{E}\Big[\|P\|_{\infty,s}^2\Big] \cdot |\dot{h}(s)| ds,$$

$$\text{For (3.20)} \Rightarrow \mathbb{E}\Big[2\int_0^T \Big|\Big\langle P(s), \Big(\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega, X(s)(\omega))\Big)\dot{h}(s)\Big\rangle\Big| ds\Big]$$
$$\leq \frac{\mathbb{E}[\|P\|_\infty^2]}{n} + n\|\varepsilon \dot{h}\|_2^2 \mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega, X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big],$$

$$\text{For (3.21)} \Rightarrow \mathbb{E}\Big[2\int_0^T \Big|\Big\langle P(s), \frac{b(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - b(s,\omega+\varepsilon h, X(s)(\omega))}{\varepsilon}\Big\rangle\Big| ds\Big]$$
$$\leq 2L \int_0^T \mathbb{E}\Big[\|P\|_{\infty,s}^2\Big] ds,$$

$$\text{For (3.22)} \Rightarrow \mathbb{E}\Big[2\int_0^T \Big|\Big\langle P(s), \frac{b(s,\omega+\varepsilon h, X(s)(\omega)) - b(s,\omega, X(s)(\omega))}{\varepsilon}\Big\rangle\Big| ds\Big]$$
$$\leq \frac{\mathbb{E}[\|P\|_\infty^2]}{n} + n\mathbb{E}\Big[\Big(\int_0^T \Big|\frac{b(s,\omega+\varepsilon h, X(s)(\omega)) - b(s,\omega, X(s)(\omega))}{\varepsilon}\Big| ds\Big)^2\Big],$$

$$\text{For (3.23)} \Rightarrow \mathbb{E}\Big[\sup_{t\in[0,T]} 2\int_0^t \Big\langle P(s), \frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X(s)(\omega))}{\varepsilon} dW(s)\Big\rangle\Big]$$
$$\leq 2C_1 \mathbb{E}\Big[\|P\|_\infty \Big(\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big)^{1/2}\Big]$$
$$\leq \frac{\mathbb{E}[\|P\|_\infty^2]}{n} + nC_1^2 \mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big]$$
$$\leq \frac{\mathbb{E}[\|P\|_\infty^2]}{n} + nC_1^2 \int_0^T \mathbb{E}\Big[\|P\|_{\infty,s}^2\Big] ds,$$

$$\text{For (3.24)} \Rightarrow \mathbb{E}\Big[\sup_{t\in[0,T]} 2\int_0^t \Big\langle P(s), \frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega, X(s)(\omega))}{\varepsilon} dW(s)\Big\rangle\Big]$$
$$\leq \frac{\mathbb{E}[\|P\|_\infty^2]}{n} + nC_1 \mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega, X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big],$$

$$\begin{aligned}\text{For (3.25)} \Rightarrow &\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega, X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big] \\ &\leq 2\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big] \\ &\quad + 2\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega, X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big],\end{aligned} \tag{3.26}$$

and finally that (3.26) $\Rightarrow 2L^2 \int_0^T \mathbb{E}\Big[\|P\|^2_{\infty,s}\Big]ds$.

Combining all these inequalities and choosing $n=6$, we can write all of this as

$$\frac{1}{6}\mathbb{E}\Big[\|P\|^2_\infty\Big] \leq \mathbb{E}\Big[\|A_\varepsilon\|^2_\infty\Big] + \bar{C}_1 \int_0^T \mathbb{E}\Big[\|P\|^2_{\infty,s}\Big]ds,$$

where $\mathbb{E}\Big[\|A_\varepsilon\|^2_\infty\Big] = O(1)$ as $\varepsilon \to 0$. Grönwall's inequality yields that $\mathbb{E}[\,\|P\|^2_\infty] = O(1)$ as $\varepsilon \to 0$.

(3.17) follow using Borel-Cantelli arguments. We have that for any $\varepsilon > 0$

$$\mathbb{P}\Big[\frac{\|X(\cdot)(\omega+\varepsilon h) - X(\cdot)(\omega)\|_\infty}{\varepsilon} > N \quad \text{i.o. } N\Big] = 0,$$

since we have

$$\sum_{N=1}^{\infty} \mathbb{P}\Big[\frac{\|X(\cdot)(\omega+\varepsilon h) - X(\cdot)(\omega)\|_\infty}{\varepsilon} > N\Big] \leq \sum_{N=1}^{\infty} \frac{\mathbb{E}\Big[\|X(\cdot)(\omega+\varepsilon h) - X(\cdot)(\omega)\|^2_\infty\Big]}{\varepsilon^2 N^2} < \infty.$$

This implies that $\|X(\cdot)(\omega+\varepsilon h) - X(\cdot)(\omega)\|_\infty/\varepsilon$ is almost surely finite. Multiplying by ε and taking the limit completes the proof.

The arguments used were not reliant on the choice of $h \in H$ but only on the square integrability of $\dot{h}$. Therefore, out statements holds true for any choice of $h \in H$.

To conclude, we show that $X(\cdot)$ satisfies Definition 3.8. A short way to demonstrate this is noting that by [MPR17], *Strong Stochastic Gâteaux Differentiability* is equivalent to *Ray Absolute Continuity* and *Stochastic Gâteaux Differentiability*. Then using standard probability results one can show that *Stochastic Gâteaux Differentiability* and Equation (3.16) imply *Strong Stochastic Gâteaux Differentiability*.

However, as we have not yet proved *Stochastic Gâteaux Differentiability* one could also argue in the following way: We have shown that the map $\varepsilon \mapsto X(\omega+\varepsilon h)$ is $\mathbb{P}$-almost surely Lipschitz continuous at $\varepsilon = 0$ for any choice of $h \in H$. To show that it is Lipschitz at any other choice of ε, we would use a Cameron Martin Transformation and show that the map $\varepsilon \mapsto X(\omega + (\varepsilon+\delta)h)$ is $\mathbb{Q}$-almost surely Lipschitz continuous then use that the probability measures $\mathbb{P}$ and $\mathbb{Q}$ are equivalent. For this step to work, we require that $\|X(\cdot)\|_\infty$ has finite second moments of order p for $p > 2$ and Proposition 3.12. $\square$

Stochastic Gateaux Differentiability of X

In order to prove the convergence in probability statement of Definition 3.9, we are going to use a combination of almost sure convergence and mean square convergence. Both of these imply convergence in probability. As an interesting aside, the speed of convergence of ε will matter here.

Theorem 3.19. *Let X be solution to the SDE* (2.1) *under Assumption 3.1. Then as $\varepsilon \to 0$*

$$\frac{X(t)(\omega+\varepsilon h) - X(t)(\omega)}{\varepsilon} \xrightarrow{\mathbb{P}} M^h(t)(\omega) = \int_0^t M_s(t)(\omega)\dot{h}(s)ds.$$

Hence X satisfies Definition 3.9, i.e. is Stochastically Gâteaux differentiable.

Proof. Methodology: We write out the SDE for $\frac{X(t)(\omega+\varepsilon h)-X(t)(\omega)}{\varepsilon} - M^h(t)(\omega)$. Then we break this up into a sequence of terms. The error terms will converge to zero in probability as we take $\varepsilon \to 0$.

We then apply Proposition 2.6 (the Grönwall type result for Convergence in Probability) to prove that the whole process will converge to zero in probability.

Firstly, we have

$$\begin{aligned}\frac{X(t)(\omega+\varepsilon h)-X(t)(\omega)}{\varepsilon}=&\int_0^t \sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h))\dot{h}(s)ds\\&+\int_0^t \Big[b(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h))-b(s,\omega,X(s)(\omega))\Big]ds\\&+\int_0^t \Big[\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h))-\sigma(s,\omega,X(s)(\omega))\Big]dW(s).\end{aligned}$$

This would mean we can write

$$\begin{aligned}&\frac{X(t)(\omega+\varepsilon h)-X(t)(\omega)}{\varepsilon}-M^h(t)(\omega)\\&=\int_0^t \Big[\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h))-\sigma(s,\omega,X(s)(\omega))\Big]\dot{h}(s)ds &(3.27)\\&+\int_0^t \Big[\frac{b(s,\omega+\varepsilon h, X(s)(\omega))-b(s,\omega,X(s)(\omega))}{\varepsilon}-\int_0^s U(r,s,\omega)\dot{h}(r)dr\Big]ds &(3.28)\\&+\int_0^t \Big[\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega))-\sigma(s,\omega,X(s)(\omega))}{\varepsilon}-\int_0^s V(r,s,\omega)\dot{h}(r)dr\Big]dW(s) &(3.29)\\&+\int_0^t \Big[\int_0^1 \nabla_x b(s,\omega+\varepsilon h,\Xi)d\xi-\nabla_x b(s,\omega,X(s)(\omega))\Big]\frac{X(s)(\omega+\varepsilon h)-X(s)(\omega)}{\varepsilon}ds &(3.30)\\&+\int_0^t \Big[\int_0^1 \nabla_x \sigma(s,\omega+\varepsilon h,\Xi)d\xi-\nabla_x \sigma(s,\omega,X(s)(\omega))\Big]\frac{X(s)(\omega+\varepsilon h)-X(s)(\omega)}{\varepsilon}dW(s) &(3.31)\\&+\int_0^t \nabla_x b(s,\omega,X(s)(\omega))\Big[\frac{X(s)(\omega+\varepsilon h)-X(s)(\omega)}{\varepsilon}-M^h(s)(\omega)\Big]ds\\&+\int_0^t \nabla_x \sigma(s,\omega,X(s)(\omega))\Big[\frac{X(s)(\omega+\varepsilon h)-X(s)(\omega)}{\varepsilon}-M^h(s)(\omega)\Big]dW(s),\end{aligned}$$

where $\Xi := X(s)(\omega)+\xi[X(s)(\omega+\varepsilon h)-X(s)(\omega)]$. Then we take $\sup$ over $t\in[0,T]$. Notice that we will not use an Itô type formula on the SDE, but proving convergence for each of the individual terms.

Firstly we consider mean convergence of (3.27),

$$\begin{aligned}&\mathbb{E}\Big[\Big(\int_0^T \Big|\Big[\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h))-\sigma(s,\omega,X(s)(\omega))\Big]\dot{h}(s)\Big|ds\Big)^2\Big]\\&\le\|\dot{h}\|_2^2\cdot\mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h))-\sigma(s,\omega,X(s)(\omega))\Big|^2ds\Big]\\&\le 2\|\dot{h}\|_2^2\Big(\mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega+\varepsilon h, X(s)(\omega))-\sigma(s,\omega,X(s)(\omega))\Big|^2ds\Big]\\&\quad+L^2T\mathbb{E}\Big[\|X(\omega+\varepsilon h)-X(\omega)\|_\infty^2\Big]\Big)\\&\le O\Big(\varepsilon^2\Big)+O\Big(\varepsilon^2\Big),\end{aligned}$$

hence this RV converges to zero in mean square as $\varepsilon\to 0$.

The term (3.28) converges in mean from Assumption 3.1 since as $\varepsilon \to 0$

$$\mathbb{E}\Big[\int_0^T \Big|\frac{b(s,\omega+\varepsilon h, X(s)(\omega)) - b(s,\omega,X(s)(\omega))}{\varepsilon} - \int_0^s U(r,s,\omega)\dot{h}(r)dr\Big|ds\Big] \to 0.$$

The term (3.29) converges in mean from Assumption 3.1, namely as $\varepsilon \to 0$

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t \Big[\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega,X(s)(\omega))}{\varepsilon} - \int_0^s V(r,s,\omega)\dot{h}(r)dr\Big]dW(s)\Big|\Big]\\
&\quad \le C_1\mathbb{E}\Big[\Big(\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega,X(s)(\omega))}{\varepsilon} - \int_0^s V(r,s,\omega)\dot{h}(r)dr\Big|^2 ds\Big)^{\frac{1}{2}}\Big] \to 0.
\end{aligned}$$

For equation (3.30), we are not able to use mean square convergence arguments because the terms $\nabla_x b(s,\omega,x)$ have polynomial growth in x and we will not necessarily have enough finite moments to ensure that this term can be dominated. Therefore, we instead prove almost sure convergence. We already have $\lim_{\varepsilon\to 0}\|X(\omega+\varepsilon h) - X(\omega)\|_\infty = 0$ almost surely. Therefore, by continuity of $\nabla_x b$ from Assumption 3.1 we get almost sure convergence of

$$\int_0^1 \nabla_x b(s,\omega+\varepsilon h,\Xi)d\xi \xrightarrow{a.s.} \nabla_x b(s,\omega,X(s)(\omega)), \quad \text{as } n\to\infty \quad \text{for almost all } s\in[0,T].$$

Since we also have almost sure finiteness of $\frac{\|X(\omega+\varepsilon h)-X(\omega)\|_\infty}{\varepsilon}$, we can conclude that (3.30) converges to zero almost surely. Therefore, it also converges in probability to zero.

(3.31) is a stochastic integral, so we will not be able to use the same almost sure convergence as for (3.30). However, since σ is Lipschitz we have $\nabla_x\sigma$ is bounded. This allows us to find a Dominating function and apply the Dominated Convergence Theorem. By the Burkholder-Davis-Gundy Inequality, we have

$$\begin{aligned}
(3.31) =&\mathbb{E}\Big[\sup_{t'\in[0,T]}\Big(\int_0^{t'}\Big[\int_0^1 \nabla_x\sigma(s,\omega+\varepsilon h,\Xi)d\xi - \nabla_x\sigma(s,\omega,X(s)(\omega))\Big]\\
&\qquad\qquad\qquad\qquad \times \frac{X(s)(\omega+\varepsilon h) - X(s)(\omega)}{\varepsilon}dW(s)\Big)^2\Big]\\
\le& C_2\mathbb{E}\Big[\int_0^T\Big|\int_0^1 \nabla_x\sigma(s,\omega+\varepsilon h,\Xi)d\xi - \nabla_x\sigma(s,\omega,X(s)(\omega))\Big|^2\cdot\Big|\frac{X(s)(\omega+\varepsilon h) - X(s)(\omega)}{\varepsilon}\Big|^2 ds\Big]\\
\le& C_2 2L^2 T\mathbb{E}\Big[\Big\|\tfrac{X(\omega+\varepsilon h)-X(\omega)}{\varepsilon}\Big\|_\infty^2\Big] < \infty.
\end{aligned}$$

Applying the Dominated Convergence Theorem, we can move the limit as $\varepsilon\to 0$ inside the integral. By continuity of $\nabla_x\sigma$, we get

$$\lim_{\varepsilon\to 0}\Big[\int_0^1 \nabla_x\sigma(s,\omega+\varepsilon h,\Xi)d\xi - \nabla_x\sigma(s,\omega,X(s)(\omega))\Big] = 0 \quad \text{a.s.}$$

which implies that (3.31) converges in mean.

We now relabel

$$U_\varepsilon^h(t)(\omega) = \tfrac{X(t)(\omega+\varepsilon h)-X(t)(\omega)}{\varepsilon} - \int_0^t M_s(t)(\omega)\dot{h}(s)ds.$$

The SDE for this process can be written as follows

$$U^h_\varepsilon(t)(\omega) = A_\varepsilon(\omega) + \int_0^t \nabla_x b(s,\omega,X(s)(\omega))U^h_\varepsilon(s)(\omega)ds + \int_0^t \nabla_x \sigma(s,\omega,X(s)(\omega))U^h_\varepsilon(s)(\omega)dW(s),$$

where the sequence A_ε is a sequence of random variables which converge in probability. Then by Proposition 2.6 the random variable $\|U^h_\varepsilon\|_\infty$ converges in probability to zero as $\varepsilon \to 0$. □

Remark 3.20. *Although convergence in probability may seem to be rather a weak result relative to the much stronger Almost sure convergence or convergence in mean square, it is actually the case that we now have both. After all, we proved that the sequence of random variables $\frac{X(\cdot)(\omega+\varepsilon h)-X(\cdot)(\omega)}{\varepsilon}$ have uniform finite p moments over ε and the limit $D^h X(\cdot)$ has finite p moments. Therefore, by standard probability theory we have mean square convergence.*

Moreover, convergence in probability implies existence of a subsequence which converges almost surely. This, combined with the Ray Absolute continuity ensures uniqueness of the limit for any *choice of subsequence which implies almost sure convergence.*

The convergence conditions on U and V in Assumption 3.1(iii) and (iv) could equivalently been stated in terms of a *Ray Absolute Continuity* and *Stochastic Gâteaux Differentiability* criterion instead of *Strong Stochastic Gâteaux Differentiability*.

Proof of Theorem 3.2

Proof of Theorem 3.2. The proof is straightforward and follows from the auxiliary results shown already. The *Ray Absolute Continuity* is proved in Proposition 3.18 while *Stochastic Gateaux Differentiability* is proved in Theorem 3.19. It is proved in [Sug85] that these conditions are equivalent to Malliavin Differentiability. Further, the Malliavin Derivative satisfies the SDE (3.1) which has a unique solution as proved in Theorem 3.17. □

3.5 Proofs of the 2nd main result - Theorem 3.7

In order to prove this result under the weakest possible conditions, we only assumed enough properties to ensure convergence. However, the Stochastic Gâteaux differentiability conditions for b and σ do not require that b and σ are Malliavin differentiable. These conditions need to be checked by the user on a case-by-case basis. Here we present an argument that under slightly stronger conditions one can establish adequate integrability and convergence of b and σ to still prove Theorem 3.2 (see Remark 3.16).

In [GS16], there is a discussion about how much continuity is required for the spacial variable in the Malliavin Derivative. The authors prove results similar to those in this paper using much weaker continuity condition, but in doing so assume the integrability of the terms $D_s b(t,\omega,X(t))$ and $D_s\sigma(t,\omega,X(t))$. In our manuscript, we were unable to ensure integrability of b and σ evaluated at X without the Lipschitz (or otherwise tractable assumptions). Weaker continuity conditions would have allowed for examples where $b(t,\omega,X(t)(\omega)$ and $\sigma(t,\omega,X(t)(\omega))$ were not adequately integrable. Therefore, for easy to check conditions, we work under Assumption 3.6 (iii') and (iv').

For simplicity, we introduce Assumption 3.21 which contains all of the relevant properties Assumption 3.6 that we require for this section. The function f is to be represents both b or σ depending on the choice of m.

Assumption 3.21. *Let $m \in \{1,2\}$. Suppose that $f : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ such that* *(i)* $\forall x \in \mathbb{R}^d$ $f(\cdot,\cdot,x) \in \mathbb{D}^{1,p}(L^m([0,T];\mathbb{R}^d))$

(ii) *f is Locally Lipschitz in the spacial variable i.e $\exists L_N > 0$ such that $\forall x, y \in \mathbb{R}^d$ such that $|x|, |y| \leq N$ and $\forall t \in [0,T]$,*

$$|f(t,\omega,x) - f(t,\omega,y)| \leq L_N|x-y| \quad \mathbb{P}\text{-almost surely.}$$

(iii) *Df are Lipschitz in their spatial variables i.e. $\exists L > 0$ constant such that $\forall (s,t) \in [0,T]^2$ and $\forall x, y \in \mathbb{R}^d$,*

$$|D_s f(t,\omega,x) - D_s f(t,\omega,y)| \leq L|x-y| \quad \mathbb{P}\text{-almost surely.}$$

Integrability and indistinguishability of the Malliavin Derivative

Lemma 3.22. *Let $m \in \{1,2\}$ and $p > 2$. Let X be solution to the SDE* (2.1) *under Assumption 2.1 and let f satisfy Assumption 3.21. Then*

$$\mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega,X(t)(\omega))|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big] < \infty.$$

Proof. By the definition of $\mathbb{D}^{1,p}(L^m([0,T];\mathbb{R}^d))$ we have for any $t \in [0,T]$

$$\mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega,0)|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big] < \infty.$$

Therefore

$$\begin{aligned}\mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t \Big|D_s f(t,\omega,X(t)(\omega))\Big|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big]&\\ \leq & 2^{\frac{p-m}{m}} \cdot 2^{\frac{p}{2}} \cdot \mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega,0)|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big]\\ &+ 2^{\frac{p-m}{m}} \cdot (2L^2)^{\frac{p}{2}} \cdot \Big(\tfrac{2}{m+2}\Big)^{\frac{p}{m}} \cdot T^{\frac{p}{m}} \cdot \mathbb{E}\Big[\|X\|_\infty^p\Big] < \infty.\end{aligned}$$

□

We have by Assumption 3.21 that for every $x \in \mathbb{R}^d$ the map $f(\cdot,\cdot,x)$ is a Malliavin differentiable process. However, it is not immediate that we have the same for $f(\cdot,\cdot,X(\cdot)(\cdot))$. We first prove an indistinguishability property for when we replace x by $X(\cdot)(\omega)$.

Lemma 3.23. *Let $m \in \{1,2\}$ and $p > 2$. Let X be solution to the SDE* (2.1) *under Assumption 2.1. Let f satisfy Assumption 3.21 and recall the directional derivative notation introduced previously, $D^h F = \langle DF, h\rangle$ for any choice of $h \in H$.*

Then, for $h \in H$ we have, (t,ω)-almost surely that

$$f\Big(t,\omega+\varepsilon h, X(t)(\omega)\Big) - f\Big(t,\omega,X(t)(\omega)\Big) = \int_0^\varepsilon D^h f(t,\omega+rh,X(t)(\omega))dr.$$

Proof. We have that $\forall x \in \mathbb{R}^d$ that $\exists C_x \subset [0,T] \times \Omega$ with $\mathbb{E}[\int_0^T \mathbb{1}_{C_x}(t,\omega)dt] = 0$, dependent on the choice of x, for which $\forall (t,\omega) \in [0,T] \times \Omega \backslash C_x$ that

$$f(t,\omega + \varepsilon h, x) - f(t,\omega,x) = \int_0^\varepsilon D^h f(t,\omega + rh, x)dr. \tag{3.32}$$

We wish to prove that we can choose a null set C which is independent of x outside of which the equality holds. To do this, we prove almost sure continuity with respect to x of both the left and right hand side of (3.32).

Almost sure continuity of the left hand side is immediate since f is locally Lipschitz. For the right hand side, we use the Lipschitz properties of the Malliavin derivative. Let r_i be an enumeration of the rationals $\mathbb{Q}^d$. Then we have $\bigcup_i C_{r_i}$ is also a null set since it is the countable union of null sets. Then for $(t,\omega) \in [0,T] \times \Omega \backslash \Big(\bigcup_i C_{r_i} \Big)$ and $\forall x \in \mathbb{Q}^d$ equation (3.32) holds. Then by the continuity of f and its Malliavin derivative we conclude that this also holds $\forall x \in \mathbb{R}^d$. □

Ray Absolute Continuity and Gâteaux Differentiability

We start by establishing Ray Absolute Continuity using Borel-Cantelli.

Lemma 3.24. *Let $m \in \{1,2\}$ and $p > 2$. Let X be solution to the SDE* (2.1) *under Assumption 2.1. Let f satisfy Assumption 3.21. Then*

$$\mathbb{E}\Big[\Big(\int_0^T \big| f(t,\omega + \varepsilon h, X(t)(\omega)) - f(t,\omega, X(t)(\omega)) \big|^m dt \Big)^{\frac{2}{m}} \Big] = O(\varepsilon^2) \quad \textit{as } \varepsilon \to 0.$$

Proof. Fix $\varepsilon > 0$. By Lemma 3.23, for almost everywhere $(t,\omega) \in [0,T] \times \Omega$ we have that

$$\big| f(t,\omega + \varepsilon h, X(t)(\omega)) - f(t,\omega, X(t)(\omega)) \big| = \Big| \int_0^\varepsilon D^h f(t,\omega + rh, X(t)(\omega))dr \Big|.$$

Arguing from this, we have with the help of the directional derivative D^h, Jensen and reverse Jensen inequality plus the growth condition implied in Assumption 3.21,

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^T \big| f(t,\omega + \varepsilon h, X(t)(\omega)) - f(t,\omega, X(t)(\omega)) \big|^m dt \Big)^{\frac{2}{m}} \Big] \\
&\qquad = \mathbb{E}\Big[\Big(\int_0^T \Big| \int_0^\varepsilon D^h f(t,\omega + rh, X(t)(\omega))dr \Big|^m dt \Big)^{\frac{2}{m}} \Big] \\
&\qquad \leq \varepsilon \int_0^\varepsilon \mathbb{E}\Big[\Big(\int_0^T |D^h f(t,\omega + rh, X(t)(\omega))|^m dt \Big)^{\frac{2}{m}} \Big] dr \\
&\qquad \leq \varepsilon \|\dot{h}\|_2^2 \int_0^\varepsilon \mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega + rh, X(t)(\omega))|^2 ds \Big)^{\frac{m}{2}} dt \Big)^{\frac{2}{m}} \Big] dr \\
&\qquad \leq \varepsilon \|\dot{h}\|_2^2 \cdot 2^{\frac{2}{m}} \int_0^\varepsilon \mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega + rh, 0)|^2 ds \Big)^{\frac{m}{2}} dt \Big)^{\frac{2}{m}} \Big] dr \qquad (3.33) \\
&\qquad\quad + \varepsilon \|\dot{h}\|_2^2 \cdot 2^{\frac{2}{m}} L^2 \Big(\frac{2}{m+2} \Big)^{\frac{2}{m}} T^{\frac{m+2}{m}} \int_0^\varepsilon \mathbb{E}\Big[\|X\|_\infty^2 \Big] dr. \qquad (3.34)
\end{aligned}$$

Firstly, we apply a Girsanov transformation (see Proposition 3.12) on term (3.33) to get

$$
\begin{aligned}
(3.33) =&\varepsilon\|\dot{h}\|_2^2\cdot 2^{\frac{2}{m}}\int_0^\varepsilon \mathbb{E}\Big[\Big(\int_0^T\Big(\int_0^t\Big|D_s f(t,\omega,0)\Big|^2 ds\Big)^{\frac{m}{2}}\mathcal{E}(r\dot{h})(t)dt\Big)^{\frac{2}{m}}\Big]dr\\
&\leq \varepsilon\|\dot{h}\|_2^2\cdot 2^{\frac{2}{m}}\int_0^\varepsilon \mathbb{E}\Big[\Big(\int_0^T\Big(\int_0^t\Big|D_s f(t,\omega,0)\Big|^2 ds\Big)^{\frac{m}{2}}dt\Big)^{\frac{2}{m}}\sup_{t\in[0,T]}|\mathcal{E}(r\dot{h})(t)|^{\frac{2}{m}}\Big]dr\\
&\leq \varepsilon\|\dot{h}\|_2^2\cdot 2^{\frac{2}{m}}\mathbb{E}\Big[\Big(\int_0^T\Big(\int_0^t\Big|D_s f(t,\omega,0)\Big|^2 ds\Big)^{\frac{m}{2}}dt\Big)^{\frac{p}{m}}\Big]^{\frac{2}{p}}\int_0^\varepsilon \mathbb{E}\Big[\sup_{t\in[0,T]}|\mathcal{E}(r\dot{h})(t)|^{\frac{2p}{m(p-2)}}\Big]^{\frac{p-2}{p}}dr.
\end{aligned}
$$

by the Hölder Inequality. Now using Proposition 3.12, we have that

$$
\varepsilon\int_0^\varepsilon \mathbb{E}\Big[\sup_{t\in[0,T]}|\mathcal{E}(r\dot{h})(t)|^{\frac{2p}{m(p-2)}}\Big]^{\frac{p-2}{p}}dr = O(\varepsilon^2)\quad \text{as } \varepsilon\to 0.
$$

Secondly, we have

$$
(3.34) \leq \varepsilon^2\|\dot{h}\|_2^2\cdot 2^{\frac{2}{m}}L^2\Big(\frac{2}{m+2}\Big)^{\frac{2}{m}}T^{\frac{m+2}{m}}\mathbb{E}\Big[\|X\|_\infty^2\Big] = O(\varepsilon^2),
$$

as $\varepsilon\to 0$ as the remaining terms are finite constants. □

The next result establishes the Stochastic Gâteaux differentiability, see Theorem 3.11.

Lemma 3.25. *Let $m\in\{1,2\}$ and $p>2$. Let X be solution to the SDE* (2.1) *under Assumption 2.1. Let f satisfy Assumption 3.21. Then for $h\in H$*

$$
\lim_{\varepsilon\to 0}\mathbb{E}\Big[\Big(\int_0^T\Big|\frac{f(t,\omega+\varepsilon h,X(t)(\omega))-f(t,\omega,X(t)(\omega))}{\varepsilon}-D^h f(t,\omega,X(t)(\omega))\Big|^m dt\Big)^{\frac{1}{m}}\Big]=0.
$$

Proof. Let $h\in H$. It is clear that both

$$
\sup_{\varepsilon\in(0,1)}\mathbb{E}\Big[\Big(\int_0^T\Big|\frac{f(t,\omega+\varepsilon h,X(t)(\omega))-f(t,\omega,X(t)(\omega))}{\varepsilon}\Big|^m dt\Big)^{\frac{1}{m}}\Big]<\infty,
$$

$$
\mathbb{E}\Big[\Big(\int_0^T\Big|D^h f(t,\omega,X(t)(\omega))\Big|^m dt\Big)^{\frac{1}{m}}\Big]<\infty.
$$

Therefore, if we can prove this for a similar Almost sure convergence statement, we can apply the Dominated Convergence Theorem to establish the conclusion.

We have that $f(\cdot,\cdot,x)\in\mathbb{D}^{1,p}(L^m([0,T];\mathbb{R}^d))$. This is equivalent to having Ray Absolute Continuity and convergence in Probability of Gâteaux Derivatives. Therefore we have that there exists a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ such that $\forall x\in\mathbb{R}^d$ $\exists C_x\subset\Omega$ with $\mathbb{P}[C_x]=0$ and $\forall\omega\in\Omega\backslash C_x$ that

$$
\lim_{n\to\infty}\int_0^T\Big|\frac{f(t,\omega+\varepsilon_n h,x)-f(t,\omega,x)}{\varepsilon_n}-D^h f(t,\omega,x)\Big|^m dt=0.
$$

By the Ray absolute continuity of $f(\cdot,\cdot,x)$, we have that this limit is not dependent on the choice of ε_n and similarly we get

$$
\lim_{\varepsilon\to 0}\int_0^T\Big|\frac{f(t,\omega+\varepsilon h,x)-f(t,\omega,x)}{\varepsilon}-D^h f(t,\omega,x)\Big|^m dt=0.
$$

We can also substitute in using that f is Malliavin differentiable to get

$$\lim_{\varepsilon\to 0}\int_0^T \Big|\frac{1}{\varepsilon}\int_0^\varepsilon D^h f(t,\omega+rh,x)dr - D^h f(t,\omega,x)\Big|^m dt = 0,$$

in a similar fashion to Lemma 3.23.

Finally, we want to show continuity on the left hand side in x. This will come from the Lipschitz property of the Malliavin derivative. We have

$$\begin{aligned}&\Big[\frac{1}{\varepsilon}\int_0^\varepsilon D^h f(t,\omega+rh,x)dr - D^h f(t,\omega,x)\Big]^m - \Big[\frac{1}{\varepsilon}\int_0^\varepsilon D^h f(t,\omega+rh,\tilde{x})dr - D^h f(t,\omega,\tilde{x})\Big]^m\\ &\leq 2K|x-\tilde{x}|\Big[\frac{1}{\varepsilon}\int_0^\varepsilon D^h f(t,\omega+rh,x)dr + D^h f(t,\omega,x) + \frac{1}{\varepsilon}\int_0^\varepsilon D^h f(t,\omega+rh,\tilde{x})dr + D^h f(t,\omega,\tilde{x})\Big].\end{aligned}$$

Integrating over $t\in[0,T]$, we now just have to show that

$$\int_0^T \Big|\frac{1}{\varepsilon}\int_0^\varepsilon D^h f(t,\omega+rh,x)dr\Big|dt \quad \text{and} \quad \int_0^T \Big|D^h f(t,\omega,x)\Big|dt,$$

are almost surely finite for all choices of $x\in\mathbb{R}^d$. The second term is immediate since we know $\mathbb{E}\big[\int_0^T |D^h f(t,\omega,x)|^2 dt\big] < \infty$. For the first, we perform a Girsanov transformation to get

$$\begin{aligned}\mathbb{E}\Big[\Big(\int_0^T \Big|\frac{1}{\varepsilon}\int_0^\varepsilon D^h f(t,\omega+rh,x)dr\Big|dt\Big)^2\Big]&\\ \leq & T\frac{1}{\varepsilon}\int_0^\varepsilon \mathbb{E}\Big[\int_0^T |D^h f(t,\omega+rh,x)|^2 dt\Big]\\ \leq & T\frac{1}{\varepsilon}\int_0^\varepsilon \mathbb{E}\Big[\|\mathcal{E}(r\cdot h)\|_\infty^q\Big]dr\,\mathbb{E}\Big[\Big(\int_0^T |D^h f(t,\omega,x)|^2 dt\Big)^p\Big]^{\frac{1}{p}} < \infty.\end{aligned}$$

Hence both of these terms will have finite second moments for any choice of $x\in\mathbb{R}^d$. Hence they are almost surely finite (Markov inequality combined with Borel-Cantelli). This means we can take $C=\cup_{x\in\mathbb{Q}^d}C_x$, which is a null set, and outside of that null set we have that

$$\lim_{\varepsilon\to 0}\int_0^T \Big|\int_0^\varepsilon D^h f(t,\omega+rh,X(t)(\omega))dr - D^h f(t,\omega,X(t)(\omega))\Big|^2 dt \to 0.$$

□

Proof of Theorem 3.7

Proof of Theorem 3.7. The difference between Assumptions 3.1 and Assumptions 3.6 is (iii') and (iv'). Here we verify that b and σ satisfying Assumption 3.6 implies Assumptions 3.1.

Lemma 3.22 implies Assumptions 3.1 (iii) is satisfied. Lemma 3.24 and Lemma 3.25 imply Assumptions 3.1 (iv) is satisfied. In this case, the identification U,V with Db and $D\sigma$ respectively is straightforward.This also means that the Existence proof in Theorem 3.17 holds so a solution to the SDE (3.2) must exist. □

4 Parametric differentiability

In this section, we study the differentiability properties of solutions of SDEs with respect to the initial variable. For a detailed exploration of the subject of Stochastic flows, see [Kun90]. The main contribution of this section is to prove similar results for SDEs with only locally Lipschitz and monotone coefficients as opposed to previous results which rely on a Lipschitz condition. Similar problems have been studied in [RS17] for Rough Differential Equations.

4.1 Gâteaux and Frechét Differentiability of monotone SDEs

We start by recalling the concept of Gâteaux and Frechét Differentiability.

Definition 4.1 (Frechét Differentiability)**.** *Let V and W be Banach spaces and let U be an open subset of V. Let $f : U \to W$. The map f is* Gâteaux differentiable *at $x \in U$ in direction $h \in V$ if the limit*

$$\lim_{\varepsilon \to 0} \frac{f(x+\varepsilon h) - f(x)}{\varepsilon} = \frac{d}{d\varepsilon} f(x+\varepsilon h),$$

exists. The limit is called the Gâteaux derivative in direction h

The map f is said to be Frechét differentiable at $x \in U$ if there exists a bounded linear operator $A : U \to V$ such that

$$\lim_{\|h\|_V \to 0} \frac{\|f(x+h) - f(x) - Ah\|_W}{\|h\|_V} = 0.$$

The linear operator A is called the Frechét derivative of f at x

Let X_θ be the solution of the SDE (2.1). We will be showing that the map $\theta \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P}) \mapsto X_\theta(\cdot) \in \mathcal{S}^p([0,T])$ is Frechét differentiable. As we will be differentiating with respect to θ for this section, we label the dependency on θ.

Assumption 4.2. *Let $b : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ satisfy Assumption 2.1 for some $p \geq 2$. Further, suppose*

(i) For almost all $(t,\omega) \in [0,T] \times \Omega$ we have the functions $\sigma(t,\omega,\cdot)$ and $b(t,\omega,\cdot)$ have partial derivatives in all directions.

(ii) For all $x \in \mathbb{R}^d$, we have that the maps

$$x \mapsto \int_0^T \Big|\nabla_x \sigma(t,\omega,x)\Big|^2 dt \quad \text{and} \quad x \mapsto \int_0^T \Big|\nabla_x b(t,\omega,x)\Big|^2 dt \quad \text{are } \mathbb{P}\text{-almost surely continuous.}$$

Note that unlike in the Malliavin Derivative case where we were unable to take a limit in the L^p norm but only in L^q norm for $q < p$, for the Frechét derivative we are able and required to prove convergence in L^p norm.

This is because for $h \in H$ we do not necessarily have $\mathbb{E}\big[\|X(\omega+h)\|_\infty^p\big] < \infty$ given only that $\mathbb{E}\big[\|X(\omega)\|_\infty^p\big] < \infty$. However, for $h \in L^p(\Omega, \mathcal{F}_0; \mathbb{R}^d)$, we do have that $\mathbb{E}\big[\|X_{\theta+h}\|_\infty^p\big] < \infty$ given that $\mathbb{E}\big[\|X_\theta\|_\infty^p\big] < \infty$. This is the reason behind the restriction $p > 2$ in Assumption 3.1.

Theorem 4.3. *Let $p \geq 2$. Let X_θ be the solution of SDE* (2.1) *under Assumption 4.2. Then the map $\theta \to X_\theta$ is Gâteaux Differentiable in direction h and the derivative is equal to $F[h]$ the solution of the SDE* (4.1)

Further, the operator $F : L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P}) \to S^p([0,T])$ is the Frechét derivative.

The proof is given after showing several intermediary results. The first result relate to Gâteaux differentiability and its properties, we address the Frechét differentiability afterwards.

Observe that for this proof once one has established Gâteaux differentiability, extending to Frechét differentiability is remarkably easy. Gâteaux differentiability is the weaker condition and is usually considered the easier property to prove.

Existence and Uniqueness for the candidate process

Theorem 4.4. *Let $p \geq 2$ and suppose Assumption 4.2 holds. Let X_θ be the solution to* (2.1)*. Let $h \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$. Then the SDE*

$$F(t)[h] = h + \int_0^t \nabla_x b\Big(s, \omega, X_\theta(s)(\omega)\Big) F(s)[h] ds + \int_0^t \nabla_x \sigma\Big(s, \omega, X_\theta(s)(\omega)\Big) F(s)[h] dW(s), \quad (4.1)$$

has a unique solution in $\mathcal{S}^p([0,T]; \mathbb{R}^d)$.

Proof. This just follows from Theorem 2.5. We simply verify that Assumption 2.4 holds:

1. $|\nabla_x \sigma| < L$ by the Lipschitz property. Therefore, clearly $\mathbb{E}\Big[\int_0^T |\nabla_x \sigma(s, \omega, X_\theta(s))|^2 ds\Big] < \infty$.

2. From the differentiability and the monotonicity property of b, we have that $\nabla_x b$ is $\mathbb{P}$-almost surely negative semidefinite[1]. Therefore, for $z \in \mathbb{R}^d$

$$z^T \int_0^T \nabla_x b(s, \omega, X_\theta(s)) ds z \leq \int_0^T L|z|^2 ds \leq LT|z|^2,$$

Hence, using the moment estimates we conclude that $\mathbb{E}\Big[\|F[h]\|_\infty^p\Big] \lesssim \|h\|^p_{L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})}$. □

Unlike with the Malliavin Derivative, the SDE (4.1) is not a general linear stochastic differential equation. As b and σ do not have dependency on θ, we do not have extra terms akin to the Malliavin derivatives Db and $D\sigma$. This means that, unlike the Malliavin Derivative, F has finite moments of all orders provided the initial condition has adequate integrability.

Proposition 4.5. *Let $p \geq 2$. Suppose Assumption 4.2. Let X_θ be the solution to* (2.1)*. The operator $F : L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P}) \to \mathcal{S}^p([0,T])$ defined by $h \mapsto F[h]$ the solution of Equation* (4.1)*, is a bounded linear operator $\|F[h]\|_{\mathcal{S}^p} \lesssim \|h\|_{L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})}$.*

Proof. We show that the operator F is a bounded linear operator. Firstly, we show that $F[0](\cdot) = 0_d$ a.s. (0_d is the $\mathbb{R}^d$-vector of zeros). Since $F[0]$ is the solution to the SDE

$$F(t)[0] = \int_0^t \nabla_x b(s, \omega, X_\theta(s)(\omega)) F(s)[0] ds + \int_0^t \nabla_x \sigma(s, \omega, X_\theta(s)(\omega)) F(s)[0] dW(s), \quad F(0)[0] = 0$$

and this SDE has a unique solution, we only need to show that $F[0](\cdot) = 0_d$ is a solution. Clearly we have $\mathbb{P}$-almost surely that

$$\int_0^t \nabla_x b(s, \omega, X_\theta(s)(\omega)) \cdot 0_d ds = 0 \quad \text{and} \quad \int_0^t \nabla_x \sigma(s, \omega, X_\theta(s)(\omega)) \cdot 0_d dW(s) = 0,$$

[1] We do not prove this fact; it is straightforward using inner products and the definition of derivative.

so this is immediate.

Let $\lambda \in \mathbb{R}$. Next we have

$$\begin{aligned}
&F[h_1](t) + \lambda F[h_2](t) \\
&\quad = h_1 + \lambda h_2 + \int_0^t \nabla_x b(s,\omega,X_\theta(s)(\omega))F[h_1](s)ds + \lambda \int_0^t \nabla_x b(s,\omega,X_\theta(s)(\omega))F[h_2](s)ds \\
&\qquad + \int_0^t \nabla_x \sigma(s,\omega,X_\theta(s)(\omega))F[h_1](s)dW(s) + \lambda \int_0^t \nabla_x \sigma(s,\omega,X_\theta(s)(\omega))F[h_2](s)dW(s), \\
&\Big(F[h_1] + \lambda F[h_2]\Big)(t) \\
&\quad = (h_1 + \lambda h_2) + \int_0^t \nabla_x b(s,\omega,X_\theta(s)(\omega))\Big(F[h_1](s) + \lambda F[h_2](s)\Big)(s)ds \\
&\qquad + \int_0^t \nabla_x \sigma(s,\omega,X_\theta(s)(\omega))\Big(F[h_1](s) + \lambda F[h_1](s)\Big)(s)dW(s),
\end{aligned}$$

which is the same as the SDE for $F[h_1 + \lambda h_2]$. Hence, by existence and uniqueness, the two must be equal up to a null set.

Finally, to prove boundedness we observe that in Theorem 4.4 we had $\|F[h]\|_{\mathcal{S}^p} \lesssim \|h\|_{L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})}$. □

Differentiability of $\theta \mapsto X_\theta$

It is immediate to prove the stochastic stability result that $\|X_{\theta+h} - X_\theta\|_\infty = O\big(\|h\|_{L^p}\big)$ $\mathbb{P}$-almost surely as $\|h\|_{L^p} \to 0$, see Theorem 2.2 with the Borel Cantelli methods of Proposition 3.18. Hence we have

$$\lim_{\|h\|_{L^p}\to 0} \|X_{\theta+h}(\omega) - X_\theta(\omega)\|_\infty \to 0, \qquad \frac{\|X_{\theta+h}(\omega) - X_\theta(\omega)\|_\infty}{\|h\|_{L^p}} \leq N(\omega),$$

where N is a square integrable random variable independent of the choice of $\|h\|_{L^p}$.

Theorem 4.6. *Let $h \in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$. Suppose we have Assumption 4.2, let X_θ be the solution of the SDE* (2.1) *and let $F(t)[h]$ be the solution to the SDE* (4.1). *Then we have*

$$\|X_{\theta+h} - X_\theta - F[h]\|_{\mathcal{S}^p} = o\Big(\|h\|_{L^p}\Big),$$

and therefore $F[h]$ is the Gâteaux derivative of X..

Proof. Let $t \in [0,T]$. Define $\Xi(\cdot) := X_\theta(\cdot) + \xi[X_{\theta+h}(\cdot) - X_\theta(\cdot)]$ and consider

$$\begin{aligned}
\frac{X_{\theta+h}(t) - X_\theta(t) - F[h](t)}{\|h\|_{L^p}} &= \frac{(\theta+h) - \theta - h}{\|h\|_{L^p}} \\
&\quad + \int_0^t \Big[\int_0^1 \nabla_x b(s,\omega,\Xi(s))d\xi - \nabla_x b(s,\omega,X_\theta(s))\Big] \cdot \Big[\tfrac{X_{\theta+h}(s) - X_\theta(s)}{\|h\|_{L^p}}\Big] ds \qquad (4.2) \\
&\quad + \int_0^t \Big[\int_0^1 \nabla_x \sigma(s,\omega,\Xi(s))d\xi - \nabla_x \sigma(s,\omega,X_\theta(s))\Big] \cdot \Big[\tfrac{X_{\theta+h}(s) - X_\theta(s)}{\|h\|_{L^p}}\Big] dW(s) \qquad (4.3) \\
&\quad + \int_0^t \nabla_x b(s,\omega,X_\theta(s)) \Big[\tfrac{X_{\theta+h}(s) - X_\theta(s) - F(s)[h](s)}{\|h\|_{L^p}}\Big] ds \\
&\quad + \int_0^t \nabla_x \sigma(s,\omega,X_\theta(s)) \Big[\tfrac{X_{\theta+h}(s) - X_\theta(s) - F(s)[h]}{\|h\|_{L^p}}\Big] dW(s).
\end{aligned}$$

Arguing the same way as in Theorem 3.19, we show that Equation (4.2) and (4.3) converge to zero in probability as $\|h\|_{L^p} \to 0$. Then we apply Proposition 2.6 to conclude that

$$\frac{\|X_{\theta+h} - X_\theta - F[h]\|_\infty}{\|h\|_{L^p}} \xrightarrow{\mathbb{P}} 0.$$

Finally, from Theorem 2.2 and Theorem 4.4 we have that

$$\frac{\mathbb{E}\Big[\|X_{\theta+h} - X_\theta\|_\infty^p\Big]}{\|h\|_{L^p}^p} = O(1), \quad \frac{\mathbb{E}\Big[\|F[h]\|_\infty^p\Big]}{\|h\|_{L^p}^p} = O(1) \qquad \text{as} \quad \|h\|_{L^p} \to 0.$$

We conclude that the random variables must converge in mean square (and equivalently $\mathcal{S}^p$). Hence we have

$$\lim_{\|h\|_{L^p}\to 0} \frac{\|X_{\theta+h} - X_\theta - F[h]\|_{\mathcal{S}^p}}{\|h\|_{L^p}} = 0,$$

which concludes the proof. □

Proof of the Frechet differentiability theorem

Proof of Theorem 4.3. In Proposition 4.5 we proved that F is a bounded linear operator and in Theorem 4.6 we proved that it satisfies Definition 4.1. □

Remark 4.7. *The Gateaux derivative of $X_\theta(t)$ with respect to θ in direction h can be expressed explicitly as*

$$F[h](t) = h \cdot \exp\left(\int_0^t \nabla_x b\Big(s, \omega, X_\theta(s)(\omega)\Big) ds - \frac{\big|\nabla_x \sigma\big(s, \omega, X_\theta(s)(\omega)\big)\big|^2}{2} ds + \int_0^t \nabla_x \sigma\Big(s, \omega, X_\theta(s)(\omega)\Big) dW(s) \right)$$

This means we can write the Frechét derivative explicitly as

$$F(t) = I \cdot \exp\left(\int_0^t \nabla_x b\Big(s, \omega, X_\theta(s)(\omega)\Big) ds - \frac{\big|\nabla_x \sigma\big(s, \omega, X_\theta(s)(\omega)\big)\big|^2}{2} ds + \int_0^t \nabla_x \sigma\Big(s, \omega, X_\theta(s)(\omega)\Big) dW(s) \right)$$

where I is the identity operator on $L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$.

4.2 Classical differentiability of SDEs

For this section, we will be studying the specific case where $\theta = x$ (a constant point in $\mathbb{R}^d$) and our perturbations are all in the constant function directions. Fix $(t, \omega) \in [0, T] \times \Omega$ and consider the map $x \in \mathbb{R}^d \mapsto X_x(t, \omega)$. We will be proving that, with probability 1 and for Lebesgue almost all $t \in [0, T]$, it is a diffeomorphism from $\mathbb{R}^d$ to $\mathbb{R}^d$. For this section, $h \in \mathbb{R}^d$ will represent some deterministic vector in Euclidean space. We will be calculating the partial derivatives in direction h.

When we prove differentiability with respect to the initial conditions, we want to prove an almost sure statement so we are not concerned with convergence in $\mathcal{S}^p$, although we will additionally obtain this. This is on contrast to the Malliavin Derivative or Frechét case where we prove a convergence in probability statement and subsequently obtain $\mathcal{S}^p$ convergence.

The Jacobian Matrix J

Definition 4.8. *Let $p \geq 2$. Let X_x be solution to the SDE* (2.1) *under Assumption 4.2 and with initial condition $X_x(0) = x \in \mathbb{R}^d$. Let I_d be the d-dimensional identity matrix. Let $J \in \mathcal{S}^p([0,T];\mathbb{R}^{d\times d})$ be the solution of the SDE*

$$J(t) = I_d + \int_0^t \nabla_x b\Big(s,\omega, X_x(s)(\omega)\Big) J(s) ds + \int_0^t \nabla_x \sigma\Big(s,\omega, X_x(s)(\omega)\Big) J(s) dW(s). \tag{4.4}$$

Notice that Equation (4.4) is the same SDE as (2.4). This means the Jacobian has an explicit solution which will be useful in Section 5 below.

Theorem 4.9. *Let $p \geq 2$. Let X_x be solution to the SDE* (2.1) *under Assumption 4.2 and with initial condition $x \in \mathbb{R}^d$. Then the SDE* (4.4) *has a unique solution in $\mathcal{S}^p$ and the map $x \mapsto X_x$ is differentiable. The derivative is almost surely equal to the solution of the Jacobian Equation, SDE* (4.4).

Differentiability of X_x

In the previous section we proved almost sure continuity of $\|X_{x+\varepsilon h} - X_x\|_\infty/\varepsilon$, we need to show that the limit as $\varepsilon \to 0$ is equal to the solution of the Jacobian SDE.

Theorem 4.10. *Let $p \geq 2$. Let X_x be solution to the SDE* (2.1) *under Assumption 4.2 and with initial condition $x \in \mathbb{R}^d$. Then we have that $\forall t \in [0,T]$*

$$\frac{X_{x+\varepsilon h}(t)(\omega) - X_x(t)(\omega)}{\varepsilon} \to h \cdot J(t)(\omega) \quad \mathbb{P}\text{-almost surely as } \varepsilon \to 0.$$

Proof. We show convergence in probability of $\frac{X_{x+\varepsilon h}(t) - X_x(t)}{\varepsilon}$ to $h \cdot J(t)$ using Proposition 2.6. Convergence in probability will imply the existence of a subsequence which converges almost sure. Finally, we will prove continuity to conclude the limit will be almost surely unique.

Writing out the SDE for the increments' process, we have

$$\begin{aligned}
&\frac{1}{\varepsilon}\big(X_{x+\varepsilon h}(t) - X_x(t)\big) - hJ(t) \\
&\qquad = \int_0^t \Big[\int_0^1 \nabla_x b(s,\omega,\Xi) d\xi - \nabla_x b(s,\omega,X_x(s))\Big]\Big[\frac{X_{x+\varepsilon h}(s) - X_x(s)}{\varepsilon}\Big] ds && (4.5)\\
&\qquad\quad + \int_0^t \Big[\int_0^1 \nabla_x \sigma(s,\omega,\Xi) d\xi - \nabla_x \sigma(s,\omega,X_x(s))\Big]\Big[\frac{X_{x+\varepsilon h}(s) - X_x(s)}{\varepsilon}\Big] dW(s) && (4.6)\\
&\qquad\quad + \int_0^t \nabla_x b(s,\omega,X_x(s))\Big[\frac{X_{x+\varepsilon h}(s) - X_x(s)}{\varepsilon} - hJ(s)\Big] ds \\
&\qquad\quad + \int_0^t \nabla_x \sigma(s,\omega,X_x(s))\Big[\frac{X_{x+\varepsilon h}(s) - X_x(s)}{\varepsilon} - hJ(s)\Big] dW(s),
\end{aligned}$$

where $\Xi = X_x(\cdot) + \xi[X_{x+\varepsilon h}(\cdot) - X_x(\cdot)]$. As with Theorem 3.19 and Theorem 4.6, we will show that the terms (4.5) and (4.6) converge in probability to 0, then use Proposition 2.6 to conclude that

$$\Big\|\frac{X_{x+\varepsilon h}(\omega)(\cdot) - X_x(\omega)(\cdot)}{\varepsilon} - hJ(\omega)(\cdot)\Big\|_\infty \xrightarrow{\mathbb{P}} 0.$$

Using the Stochastic Stability from Theorem 2.2, we have that

$$\Big\|\frac{X_{x+\varepsilon h}(\omega)(\cdot) - X_x(\omega)(\cdot)}{\varepsilon}\Big\|_\infty < N(\omega),$$

where N is square integrable and independent on the choice of ε. $N(\omega)$ will be $\mathbb{P}$-almost surely finite so we can conclude that $\frac{\|X_{x+\varepsilon h}(\omega)(\cdot)-X_x(\omega)(\cdot)\|_\infty}{\varepsilon}$ is continuous in ε around zero.

This completes the proof. $\square$

Invertibility of the Jacobian Matrix

Next, we wish to show that the Jacobian Matrix $J(t)$ is $\mathbb{P}$-almost surely invertible for any choice of $t \in [0,T]$. Notice that due to the initial condition, we have that this is true for $t = 0$ since $J(0) = I_d$.

To prove the Jacobian is invertible, we consider a matrix valued stochastic process and observe that for any choice of $t \in [0,T]$, this process will take value equal to the left inverse of J. This proof follows that of Nualart, [Nua06, Chapter 2.3; Equation 2.8].

We introduce the SDE

$$\begin{aligned}K(t) = I_d - \int_0^t K(s)\Big[\nabla_x b(s,\omega,X(s)) - \Big\langle \nabla_x\sigma, \nabla_x\sigma\Big\rangle_{\mathbb{R}^m}\big(s,\omega,X(s)\big)\Big]ds\\ - \int_0^t K(s)\nabla_x\sigma(s,\omega,X(s))dW(s). \qquad (4.7)\end{aligned}$$

Proposition 4.11. *Let $p \geq 2$. Let X_x be solution to the SDE (2.1) under Assumption 4.2 and with initial condition $x \in \mathbb{R}^d$. Then we have the following identity $K(t)J(t) = I_d$ for all $t \in [0,T]$ $\mathbb{P}$-a.s.*

Proof. The proof of this result follows from Itô's formula. Note that as we are dealing with matrix valued processes we can not necessarily assume commutativity and we need to take care. Itô's formula for matrices gives

$$\begin{aligned}d(KJ)(t) =&K(t)dJ(t) + dK(t)J(t) + d[K,J](t)\\ =&K(t)\nabla_x b(t,\omega,X(t))J(t)dt + K(t)\sigma(t,\omega,X(t))J(t)dW(t)\\ &- K(t)\nabla_x b(t,\omega,X(t))J(t)dt - K(t)\sigma(t,\omega,X(t))J(t)dW(t)\\ &+ K(t)\Big\langle \nabla_x\sigma, \nabla_x\sigma\Big\rangle_{\mathbb{R}^m}\big(s,\omega,X(s)\big)J(t)dt\\ &- K(t)\Big\langle \nabla_x\sigma, \nabla_x\sigma\Big\rangle_{\mathbb{R}^m}\big(s,\omega,X(s)\big)J(t)dt = 0dt + 0dW(t).\end{aligned}$$

$\square$

Observe that SDE (4.7) does not necessarily satisfy Assumption 2.4 as the term $-z^T\nabla_x b(t,\omega,X(t))z$ will not be bounded above by a constant almost surely for any choice of vector $|z| = 1$. However, an explicit solution to the SDE can be written out pathwise, even if it does not have finite moments. This definition will have the property that it is the left inverse of J.

Proposition 4.12. *The determinant of the Matrix $J(t)$, denoted $D(t)$, is called the Stochastic Wronskian and satisfies the SDE*

$$\begin{aligned}dD(t) =& Tr\Big(\nabla_x b(t,\omega,X(t))\Big)D(t)dt + Tr\Big(\nabla_x\sigma(t,\omega,X(t))\Big)D(t)dW(t) \qquad (4.8)\\ &+ \Big[\Big\langle Tr\big(\nabla_x\sigma(t,\omega,X(t))\big), Tr\big(\nabla_x\sigma(t,\omega,X(t))\big)\Big\rangle_{\mathbb{R}^m} - Tr\Big(\big\langle \nabla_x\sigma(t,\omega,X(t)), \nabla_x\sigma(t,\omega,X(t))\big\rangle_{\mathbb{R}^m}\Big)\Big]D(t)dt\end{aligned}$$

with $D(0) = 1$. $D(t)$ has explicit form

$$D(t) = \exp\Bigg(\int_0^t Tr\Big(\nabla_x b(t,\omega,X(t))\Big) - \frac{1}{2} Tr\Big(\big\langle \nabla_x \sigma(t,\omega,X(t)), \nabla_x \sigma(t,\omega,X(t))\big\rangle_{\mathbb{R}^m}\Big) \tag{4.9}$$
$$+ \int_0^t Tr\Big(\nabla_x \sigma(t,\omega,X(t))\Big) dW(t)\Bigg).$$

Proof. The proof can be found in [Mao08, Theorem 3.2.2]. The proof involves applying Itô's formula to the determinant of $J(t)$ and establishing that it satisfies Equation (4.8). Then one applies Itô's formula to Equation (4.9) and verifies that this likewise satisfies (4.8). Finally, by Theorem 2.5, the solution is unique. □

Therefore, we can conclude the $D(t)$ is almost surely positive and therefore the process K is $\mathbb{P}$-almost surely the inverse (left or right) of J.

5 Applications

In this section, we recover and discuss some standard applications of Malliavin Differentiation and evaluate some of the problems that occur under our framework.

Representation formulae

Firstly, we present a way of writing the Malliavin Derivative of X_θ in terms of the Jacobian.

Proposition 5.1 (Representation formulae)**.** *Let X_x be solution to the SDE* (2.1) *under Assumption 3.1 and with initial condition $X_x(0) = x \in \mathbb{R}^d$. Let J satisfy the SDE* (4.4)*. Consider the SDE for the process $J(t)J(s)^{-1}$ for $t > s$.*

$$\begin{aligned} J_s(t) =& J(t)J(s)^{-1} \\ =& J(s)J(s)^{-1} + \int_s^t \nabla_x b(r,\omega,X(r)(\omega)) J(r)J(s)^{-1} dr + \int_s^t \nabla_x \sigma(r,\omega,X(r)(\omega)) J(r)J(s)^{-1} dW(r) \\ =& I_d + \int_s^t \nabla_x b(r,\omega,X(r)(\omega)) J_s(r) dr + \int_s^t \nabla_x \sigma(r,\omega,X(r)(\omega)) J_s(r) dW(r). \end{aligned} \tag{5.1}$$

Equation (5.1) *is the Fundamental Matrix of the Linear Stochastic Differential Equation* (3.1)*. As such, under Assumption 3.1 the Malliavin Derivative of X can be expressed for $t > s$ as*

$$D_s X(t) = J_s(t) A(s,t),$$

where $A(s,t)$ is defined for $t > s$ as

$$A(s,t) = \sigma(s,\omega,X(s)(\omega)) + \int_s^t J_s(r)^{-1}\Big(U(s,r,\omega) - \Big\langle \nabla_x \sigma(r,\omega,X(r)(\omega)), V(s,r,\omega)\Big\rangle_{\mathbb{R}^m}\Big) dr$$
$$+ \int_s^t J_s^{-1}(r) V(s,r,\omega) dW(r).$$

Proof. The proof of this representation formula follows the same ideas as Theorem 3.17. Equation (3.1) is an SPDE, so we project from the infinite dimensional space into a finite dimensional space. We follow the method of [Mao08, Theorem 3.3.1] to solve the solution explicitly in the projection space then use the Dominated Convergence Theorem to ensure the passage to the limit.

□

Absolute Continuity

In [Nua06, Theorem 2.3.1], it is proved that the solution of a Stochastic Differential Equation with Lipschitz, deterministic coefficients and elliptic diffusion term has a law which is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$. This proof can be easily extended to the case where the drift term has monotone growth.

However, the proof is more involved if we allow the coefficients to be random.

Theorem 5.2. *Let X_x be solution to the SDE* (2.1) *under Assumption 3.1 and with initial condition $X_x(0) = x \in \mathbb{R}^d$. Suppose additionally that $\forall z \in \mathbb{R}^d$ that*

$$z^T A(s,t)A(s,t)^T z > \lambda(s,t)|z|^2 \geq 0, \qquad \int_0^t \lambda(s,t)ds > 0 \quad \mathbb{P}\text{-almost surely.}$$

Then the law of $X_x(t)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$.

Proof. For this proof, we show that the Malliavin matrix is $\mathbb{P}$-almost surley non zero and recall [Nua06, Corollary 2.1.2].

The Malliavin Matrix, $Q(t)$ is defined to be

$$Q(t) = \int_0^t D_s X(t) D_s X(t)^T ds = J(t) \int_0^t K(s)A(s,t)A(s,t)^T K(s)^T ds J(t)^T$$

Therefore, for $z \in \mathbb{R}^d$ we have $z^T Q(t) z \geq \int_0^t \lambda(s,t)|K(s)|^2 ds \cdot |J(t)|^2 \cdot |z|^2$ which is greater than zero because $|J|, |K| > 0$ and by assumption. □

Remark 5.3. *Observe that the Ellipticity condition for σ is no longer enough to ensure that the law is absolutely continuous. When b and σ are deterministic, U and V are uniformly 0 and Ellipticity is enough.*

Bismut-Elworthy-Li formula

In [Elw92], the author uses Malliavin Differentiability of an SDE X_x to prove differentiability for functions of the form

$$u(x) = \mathbb{E}[\phi(X_x(t))]$$

where ϕ is assumed to be a continuous function and $t \in [0,T]$. This was later extended in [FLL+99] and [FLLL01] to cover functions ϕ which are integrable and even measurable (provided u remains finite).

Define for $t \in (0,T]$ the set $\Gamma_t = \big\{a \in L^2([0,T]); \int_0^t a(s)ds = 1\big\}$.

Theorem 5.4 (Bismut-Elworthy-Li formula)**.** *Let $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ be a bounded, measurable function. Let X_x be solution to the SDE* (2.1) *under Assumption 3.1 and with initial condition $X_x(0) = x \in \mathbb{R}^d$. Let $t \in (0,T]$. Suppose additionally that*

1. *$\forall s \in [0,t]$ the matrix $A(s,t)$ has a right inverse.*
2. *$\exists a \in \Gamma_t$ such that $a(\cdot)A(\cdot,t)^{-1}J(\cdot) \in dom(\delta)$.*

Then

$$\nabla_x \mathbb{E}\Big[\Phi(X_x(t)\Big] = \mathbb{E}\Big[\Phi(X_x(t))\delta\Big(a(s)A(s,t)^{-1}J(s)\Big)\Big].$$

Proof. We give only a sketch of the proof. For a more detailed proof, see [FLL+99] and [FLLL01]. First suppose that Φ is continuously differentiable with bounded derivatives, then

$$\nabla_x \mathbb{E}\Big[\Phi(X_x(t))\Big] = \mathbb{E}\Big[\nabla_x \Phi(X_x(t))\Big] = \mathbb{E}\Big[\nabla \Phi(X_x(t)) J(t)\Big] = \mathbb{E}\Big[\nabla \Phi(X_x(t)) D_s X_x(t) A(s,t)^{-1} J(s)\Big].$$

Multiplying both sides by $a \in \Gamma_s$, integrating over $[0,t]$ (using $\int_0^t a(s)ds = 1$ on the LHS) and Fubini gives

$$\begin{aligned}
\nabla_x \mathbb{E}\Big[\Phi(X_x(t))\Big] =& \mathbb{E}\Big[\int_0^t a(s) \nabla \Phi(X_x(t)) D_s X_x(t) A(s,t)^{-1} J(s) ds\Big] \\
=& \mathbb{E}\Big[\int_0^t D_s\Big(\Phi(X_x(t))\Big) a(s) A(s,t)^{-1} J(s) ds\Big] = \mathbb{E}\Big[\Phi\big(X_x(t)\big)\delta\Big(a(s)A(s,t)^{-1}J(s)\Big)\Big],
\end{aligned}$$

where in the last line we used integration-by-parts formula.

Secondly, let Φ be bounded and measurable. Then using that C_b^1 is dense in the set of bounded measurable functions, we approximate Φ by a sequence of functions $\Phi_n \in C_b^1$. Finally, using a domination argument it is shown that one can swap the limits and integrals and one reaches the conclusion. $\square$

A Proofs

A.1 The existence and uniqueness theorem plus moment calculations

Proof of Theorem 2.2. As $p \geq 2$, we also have that

$$\mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega,X_\theta(s))\Big|^2 ds\Big] \leq 2\mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega,0)\Big|^2 ds\Big] + 2L^2 T \mathbb{E}\Big[\|X\|_\infty^2\Big] < \infty.$$

This means we can use [Øks03, Theorem 3.2.5] to get $\mathbb{P}$-almost sure continuity of the stochastic integral. The drift term is a Lebesgue integral so likewise is continuous in time. Hence $\mathbb{P}$-almost sure continuity of $t \mapsto X_\theta(t)$ is immediate.

Finally, let $t \in [0,T]$ and $\xi, \theta \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$. We have

$$\begin{aligned}
X_\xi(t) - X_\theta(t) = \xi - \theta + \int_0^t \Big[b(s,\omega,X_\xi(s)) - b(s,\omega,X_\theta(s))\Big] ds \\
+ \int_0^t \Big[\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\Big] dW(s).
\end{aligned}$$

We write $Q(s) = X_\xi(s) - X_\theta(s)$ and by applying Itô s formula with $f(x) = |x|^p$ we get

$$\begin{aligned}
|Q(t)|^p = |\xi - \theta|^p &+ p \int_0^t |Q(s)|^{p-2} \Big\langle Q(s), b(s,\omega,X_\xi(s)) - b(s,\omega,X_\theta(s))\Big\rangle ds \\
&+ p \int_0^t |Q(s)|^{p-2} \Big\langle Q(s), \Big[\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\Big] dW(s)\Big\rangle \\
&+ \frac{p}{2} \int_0^t |Q(s)|^{p-2} \cdot \Big|\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\Big|^2 ds \\
&+ \frac{p(p-2)}{2} \int_0^t |Q(s)|^{p-4} \Big|Q(s)^T \cdot \big[\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\big]\Big|^2 ds.
\end{aligned}$$

Taking a supremum over time and then taking expectations, we get

$$\mathbb{E}\Big[\|X_\xi - X_\theta\|_\infty^p\Big] = \mathbb{E}\Big[\|Q\|_\infty^p\Big] \le \mathbb{E}\Big[|\xi-\theta|^p\Big]$$

$$+ p\mathbb{E}\Big[\int_0^T |Q(s)|^{p-2}\Big|\Big\langle Q(s), b(s,\omega,X_\xi(s)) - b(s,\omega,X_\theta(s))\Big\rangle\Big| ds\Big] \tag{A.1}$$

$$+ p\mathbb{E}\Big[\sup_{t\in[0,T]}\int_0^t |Q(s)|^{p-2}\Big\langle Q(s), \Big[\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\Big]dW(s)\Big\rangle\Big] \tag{A.2}$$

$$+ \frac{p}{2}\mathbb{E}\Big[\int_0^T |Q(s)|^{p-2}\Big|\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\Big|^2 ds\Big] \tag{A.3}$$

$$+ \frac{p(p-2)}{2}\mathbb{E}\Big[\int_0^T |Q(s)|^{p-4}|Q(s)^T\Big[\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\Big]|^2 ds\Big]. \tag{A.4}$$

Firstly by monotonicity of b we have (A.1) $\le pL\int_0^T \mathbb{E}\Big[\|Q\|_{\infty,s}^p\Big]ds$. Secondly, by the Burkholder-Davis-Gundy inequality we have

$$\text{(A.2)} \le pC_1L\mathbb{E}\Big[\|Q\|_\infty^{\frac{p}{2}}\Big(\int_0^T \|Q\|_{\infty,s}^p ds\Big)^{\frac{1}{2}}\Big] \le \frac{\mathbb{E}\Big[\|Q\|_\infty^p\Big]}{2} + \frac{p^2C_1^2L^2}{2}\int_0^T \mathbb{E}\Big[\|Q\|_{\infty,s}^p\Big]ds.$$

Finally, we have

$$\text{(A.3)} \le \frac{pL}{2}\int_0^T \mathbb{E}\Big[\|Q\|_{\infty,s}^p\Big]ds \qquad \text{and} \qquad \text{(A.4)} \le \frac{p(p-2)L}{2}\int_0^T \mathbb{E}\Big[\|Q\|_{\infty,s}^p\Big]ds.$$

Hence we have

$$\frac{1}{2}\mathbb{E}\Big[\|X_\xi - X_\theta\|_\infty^p\Big] \le \mathbb{E}\Big[|\xi-\theta|^p\Big] + \widehat{C}\int_0^T \mathbb{E}\Big[\|X_\xi - X_\theta\|_{\infty,s}^p\Big]ds,$$

where $\widehat{C} = \frac{p^2L(LC_1^2+2)}{2}$. Applying Grönwall's inequality gives that $\mathbb{E}\Big[\|X_\xi - X_\theta\|_\infty^p\Big] \lesssim \mathbb{E}[|\xi-\theta|^p]$. $\square$

Moment Calculations for Theorem 2.5 . Fix $t \in [0,T]$ and using Itô's formula with $f(x) = |x|^p$ and X_θ satisfying Equation (2.3), we get that

$$\begin{aligned}
|X_\theta(t)|^p = |\theta|^p &+ p\int_0^t |X_\theta(s)|^{p-2}\langle X_\theta(s), B(s,\omega)X_\theta(s)\rangle ds + p\int_0^t |X_\theta(s)|^{p-2}\langle X_\theta(s), b(s,\omega)\rangle ds \\
&+ p\int_0^t |X_\theta(s)|^{p-2}\langle X_\theta(s), \Sigma(s,\omega)X_\theta(s)dW(s)\rangle + p\int_0^t |X_\theta(s)|^{p-2}\langle X_\theta(s), \sigma(s,\omega)dW(s)\rangle \\
&+ \frac{p}{2}\int_0^t |X_\theta(s)|^{p-2}\Big[\Sigma(s,\omega)X_\theta(s) + \sigma(s,\omega)\Big]^2 ds \\
&+ \frac{p(p-2)}{2}\int_0^t |X_\theta(s)|^{p-4}\Big\langle X_\theta(s), \big[\Sigma(s,\omega)X_\theta(s) + \sigma(s,\omega)\big]\Big\rangle^2 ds.
\end{aligned}$$

Take a supremum over $t \in [0,T]$ and expectations yields

$$\begin{aligned}\mathbb{E}\Big[\|X_\theta\|^p\Big] \leq & \mathbb{E}\Big[|\theta|^p\Big] \\ & + p\mathbb{E}\Big[\int_0^T |X_\theta(s)|^{p-2}\Big\langle X_\theta(s), \big[B(s,\omega)X_\theta(s) + b(s,\omega)\big]\Big\rangle ds\Big] \qquad \text{(A.5)} \\ & + p\mathbb{E}\Big[\sup_{t\in[0,T]}\int_0^t |X_\theta(s)|^{p-2}\Big\langle X_\theta(s), \big[\Sigma(s,\omega)X_\theta(s) + \sigma(s,\omega)\big]dW(s)\Big\rangle\Big] \qquad \text{(A.6)} \\ & + \frac{p(p-1)}{2}\mathbb{E}\Big[\int_0^T |X_\theta(s)|^{p-2}\Big[\Sigma(s,\omega)X_\theta(s) + \sigma(s,\omega)\Big]^2 ds\Big].\end{aligned}$$

Fix $n \in \mathbb{N}$ to be chosen later. Throughout the next three arguments, we use Young's Inequality. Using the negative semidefinite properties of B, we get that

$$\text{(A.5)} \leq pL\int_0^T \mathbb{E}\Big[\|X_\theta\|^p_{\infty,s}\Big]ds + \frac{\mathbb{E}\Big[\|X_\theta\|^p_\infty\Big]}{n} + n^{p-1}(p-1)^{p-1} \times \mathbb{E}\Big[\Big(\int_0^T |b(s,\omega)|ds\Big)^p\Big].$$

Secondly, using the Burkholder-Davis-Gundy Inequality gives that

$$\begin{aligned}\text{(A.6)} \leq & \frac{2\mathbb{E}\Big[\|X_\theta\|^p_\infty\Big]}{n} + \frac{p^2C_1^2 n\sqrt{2}}{4}\int_0^T \mathbb{E}\Big[\|X_\theta\|^p_{\infty,s}\Big]\|\Sigma(s,\cdot)\|^2_{L^\infty}ds \\ & + C_1^p n^{p-1} p^{\frac{p}{2}}(p-2)^{\frac{p-2}{2}}\frac{2^{\frac{p}{4}}}{2^{p-1}}\cdot\mathbb{E}\Big[\Big(\int_0^T |\sigma(s,\omega)|^2 ds\Big)^{\frac{p}{2}}\Big].\end{aligned}$$

Thirdly, we have

$$\begin{aligned}\text{(A.6)} \leq & \frac{\mathbb{E}\Big[\|X_\theta\|^p_\infty\Big]}{n} + p(p-1)\int_0^T \mathbb{E}\Big[\|X_\theta\|^p_{\infty,s}\Big]\|\Sigma(s,\cdot)\|_{L^\infty}ds \\ & + 2[n(p-2)]^{\frac{p-2}{2}}(p-1)^{\frac{p}{2}}\mathbb{E}\Big[\Big(\int_0^T |\sigma(s,\omega)|^2 ds\Big)^{\frac{p}{2}}\Big].\end{aligned}$$

When applying Young's Inequality for the case $p = 2$, we use the convention that $0^0 = 1$. Adding these together, we have that there are constants $\widetilde{C_1}, \widetilde{C_2}$ and $\widetilde{C_3}$ such that

$$\begin{aligned}\frac{1}{5}\mathbb{E}\Big[\|X_\theta\|^p_\infty\Big] \leq & \mathbb{E}\Big[|\theta|^p\Big] + \widetilde{C_1}\mathbb{E}\Big[\Big(\int_0^T |b(s,\omega)|ds\Big)^p\Big] + \widetilde{C_2}\mathbb{E}\Big[\Big(\int_0^T |\sigma(s,\omega)|^2 ds\Big)^{\frac{p}{2}}\Big] \\ & + \widetilde{C_3}\int_0^T \big[1 + \|\Sigma(s,\cdot)\|_{L^\infty}\big]\mathbb{E}\Big[\|X_\theta\|^p_{\infty,s}\Big]ds.\end{aligned}$$

Applying Grönwall Inequality yields

$$\begin{aligned}\mathbb{E}\Big[\|X_\theta\|^p_\infty\Big] \leq 5 & \left(\mathbb{E}\Big[|\theta|^p\Big] + \widetilde{C_1}\mathbb{E}\Big[\Big(\int_0^T |b(s,\omega)|ds\Big)^p\Big] + \widetilde{C_2}\mathbb{E}\Big[\Big(\int_0^T |\sigma(s,\omega)|^2 ds\Big)^{\frac{p}{2}}\Big]\right) \\ & \times \exp\Big(5\widetilde{C_3}\int_0^T \big[1 + \|\Sigma(s,\cdot)\|_{L^\infty}\big]ds\Big).\end{aligned}$$

□

A.2 Proof of Proposition 3.12

Proof of Proposition 3.12. Let $t \in [0,T]$ and $h \in H$. The process $\mathcal{E}(h)(\cdot)$ satisfies the SDE

$$\mathcal{E}(h)(t) = 1 + \int_0^t \mathcal{E}(h)(s)\dot{h}(s)dW(s).$$

Therefore, using Itô's formula we have

$$|\mathcal{E}(h)(t)|^p = 1 + p\int_0^t |\mathcal{E}(h)(s)|^p \dot{h}(s)dW(s) + \frac{p(p-1)}{2}\int_0^t |\mathcal{E}(h)(s)|^p |\dot{h}(s)|^2 ds.$$

We take supremum over time and expectations to get

$$\mathbb{E}\Big[\|\mathcal{E}(h)(\cdot)\|_\infty^p\Big] \leq 1 + p\mathbb{E}\Big[\sup_{t\in[0,T]}\int_0^t |\mathcal{E}(h)(s)|^p \dot{h}(s)dW(s)\Big] + \frac{p(p-1)}{2}\mathbb{E}\Big[\int_0^T |\mathcal{E}(h)(s)|^p |\dot{h}(s)|^2 ds\Big].$$

We have that

$$\mathbb{E}\Big[\int_0^T |\mathcal{E}(h)(s)|^p |\dot{h}(s)|^2 ds\Big] \leq \int_0^T \mathbb{E}\Big[\|\mathcal{E}(h)(\cdot)\|_{\infty,s}^p\Big]|\dot{h}(s)|^2 ds.$$

Similarly using the Burkholder-Davis-Gundy Inequality

$$\begin{aligned} p\mathbb{E}\Big[\sup_{t\in[0,T]}\int_0^t |\mathcal{E}(h)(s)|^p \dot{h}(s)dW(s)\Big] \leq & pC_1\mathbb{E}\Big[\Big(\int_0^T |\mathcal{E}(h)(s)|^{2p}|\dot{h}(s)|^2 ds\Big)^{\frac{1}{2}}\Big] \\ \leq & \frac{\mathbb{E}\Big[\|\mathcal{E}(h)(\cdot)\|_\infty^p\Big]}{2} + \frac{p^2C_1^2}{2}\int_0^T \mathbb{E}\Big[\|\mathcal{E}(h)(\cdot)\|_{\infty,s}^p\Big]|\dot{h}(s)|^2 ds. \end{aligned}$$

Adding these together gives

$$\frac{\mathbb{E}\Big[\|\mathcal{E}(h)(\cdot)\|_\infty^p\Big]}{2} \leq 1 + \int_0^T \mathbb{E}\Big[\|\mathcal{E}(h)(\cdot)\|_{\infty,s}^p\Big]\Big(\frac{p(p-1)}{2} + \frac{p^2C_1^2}{2}\Big)|\dot{h}(s)|^2 ds.$$

Applying Grönwalls inequality and using that $h \in H$ and hence $\dot{h} \in L^2([0,T])$ we conclude that

$$\mathbb{E}\Big[\|\mathcal{E}(h)(\cdot)\|_\infty^p\Big] \leq 2\exp\Big((p^2(1+C_1^2) - p)\int_0^T |\dot{h}(s)|^2 ds\Big) < \infty.$$

□

References

[Elw92] K. D. Elworthy, *Stochastic flows on Riemannian manifolds*, Diffusion processes and related problems in analysis, volume ii: Stochastic flows, 1992, pp. 37–72.

[FLLL01] E. Fournié, J.-M. Lasry, J. Lebuchoux, and P.-L. Lions, *Applications of Malliavin calculus to Monte-Carlo methods in finance. II*, Finance and Stochastics **5** (2001Apr), no. 2, 201–236.

[FLL+99] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, and N. Touzi, *Applications of Malliavin calculus to Monte-Carlo methods in finance*, Finance and Stochastics **3** (1999Aug), no. 4, 391–412.

[GS16] C. Geiss and A. Steinicke, *Malliavin derivative of random functions and applications to Lévy driven BSDEs*, Electronic Journal of Probability **21** (2016).

[HJK11] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden, *Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **467** (2011), no. 2130, 1563–1576.

[IMPR16] P. Imkeller, T. Mastrolia, D. Possamaï, and A. Réveillac, *A note on the Malliavin-Sobolev spaces*, Statist. Probab. Lett. **109** (2016), 45–53. MR3434959

[Øks03] B. Øksendal, *Stochastic differential equations: An introduction with applications*, Hochschultext / Universitext, Springer, 2003.

[Kun90] H. Kunita, *Stochastic flows and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, vol. 24, Cambridge University Press, Cambridge, 1990. MR1070361 (91m:60107)

[Mao08] X. Mao, *Stochastic differential equations and applications*, Horwood, 2008.

[MPR17] T. Mastrolia, D. Possamaï, and A. Réveillac, *On the Malliavin differentiability of BSDEs*, Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), no. 1, 464–492. MR3606749

[MSS12] C. J. Mode, C. K. Sleeman, and W. Scientific, *Stochastic processes in genetics and evolution: Computer experiments in the quantification of mutation and selection*, World Scientific Publishing Company, 2012.

[Nua06] D. Nualart, *The Malliavin calculus and related topics*, Vol. 1995, Springer, 2006.

[RS17] S. Riedel and M. Scheutzow, *Rough differential equations with unbounded drift term*, Journal of Differential Equations **262** (2017), no. 1, 283 –312.

[Sug85] H. Sugita, *On a characterization of the Sobolev spaces over an abstract Wiener space*, J. Math. Kyoto Univ. **25** (1985), no. 4, 717–725. MR810975

[TZ13] M. Tahmasebi and S. Zamani, *Weak differentiability of solutions to SDEs with semi-monotone drifts*, arXiv preprint arXiv:1309.0619 (2013).

[ÜZ00] A. S. Üstünel and M. Zakai, *Transformation of measure on Wiener space*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000. MR1736980